Applied Energy 427 (2027) 128871

Contents lists available at ScienceDirect

Applied Energy

journal homepage: www.elsevier.com/locate/apenergy

Compound deployment risk analysis in direct air capture, green ammonia and long-duration energy storage

Sinan Küfeoğlu

School of Technology and Innovations, University of Vaasa, Vaasa 65200, Finland

HIGHLIGHTS

- A compound deployment risk index scores DAC, green ammonia and LDES on eight causes.
- Index orderings are invariant from fully compensatory to strictly conjunctive aggregation.
- Modelled DAC removal cost exceeds all compliance anchors and meets only the top voluntary tier.
- Breakeven compliance penalties exceed IMO remedial-unit prices except at best-resourced sites.
- Great Britain's storage scheme drew 52.6 GW of bids and shortlisted mostly incumbent technologies.

ARTICLE INFO

Keywords:
Direct air capture
Green ammonia
Long-duration energy storage
Technology deployment
Compound risk
Demand-side policy

ABSTRACT

Direct air capture (DAC), green ammonia and long-duration energy storage (LDES) projects have attracted large public commitments, yet project cancellations are mounting. This paper argues that deployment fails when several barriers bind at once, chiefly absent demand, uncompetitive cost, policy uncertainty and unbankable finance, and that a technology's prospects are set by its weakest barrier rather than its average strength. That logic is formalised in a Compound Deployment Risk Index, stress-tested for robustness, and paired with a levelised-cost model for each technology, using sourced data through mid-2026. The results show why each technology stalls. Capturing CO2 from air costs 480–1000 USD per tonne, more than any compliance scheme currently pays. Green ammonia becomes competitive as a shipping fuel only at a carbon penalty of 370 to 680 USD per tonne, above the 100 to 380 USD proposed by the International Maritime Organization. Furthermore, in Great Britain's storage auction, new battery chemistries were undercut by lithium-ion and pumped hydro at every procured duration. The policy conclusion is that creating demand is necessary but not sufficient: it must be paired with cost reduction and de-risked finance for a self-sustaining industry to emerge.

1. Introduction

Direct air capture (DAC), green ammonia, and long-duration energy storage (LDES) are poised for a similar moment in 2025 to 2026: rapid growth in policy commitments, accelerating project announcements, and optimistic learning-curve projections, accompanied by a rising count of cancellations and stalled final investment decisions (FIDs). For capital-intensive climate technologies, laboratory readiness and demonstrated engineering performance no longer predict commercial uptake; what predicts it is whether a set of structural market, enabling and execution conditions are simultaneously in place when capital is committed. When several fail at once, they amplify one another, and the resulting compound risk is systematically underestimated by single-cause project appraisal.

This pattern echoes a long-standing debate in the economics of technological change. Jaffe, Newell and Stavins [1] show that market failures associated with environmental externalities compound with market failures in the innovation and diffusion of new technology, so that environmentally beneficial technologies are systematically under-

Abbreviations: CDRI, Compound Deployment Risk Index; CRI, Commercial Readiness Index; DAC, direct air capture; LDES, long-duration energy storage; LCOD, levelised cost of removal; LCOH, levelised cost of hydrogen; LCOS, levelised cost of storage; FID, final investment decision; IMO, International Maritime Organization; NZF, Net-Zero Framework; MGOe, marine-gas-oil equivalent; TRL, Technology Readiness Level; WACC, weighted average cost of capital; RTE, round-trip efficiency; DOE, (U.S.) Department of Energy; DESNZ, (UK) Department for Energy Security and Net Zero; PSH, pumped storage hydro; CAES, compressed air energy storage; VRFB, vanadium redox flow battery; NZE, Net Zero Emissions (by 2050 Scenario); SEJ, structured expert judgment..

E-mail address: sinan.kufeoglu@uwasa.fi.

https://doi.org/10.1016/j.apenergy.2026.128871
Received 8 July 2026; Received in revised form 3 September 2026; Accepted 14 September 2026
Available online 27 September 2026

Nomenclature

ak	activation score of cause k (dimensionless, 0.1 to 0.9)
pk	pass score of cause k (= 1 minus ak)
V	compound viability
g	readiness score (geometric mean)
n	number of causes ($n = 8$)
ρ	pairwise rank correlation imposed among cause-level coding perturbations
A	DAC non-energy annualised cost (USD/t CO2)
etaE	DAC electricity intensity (MWh/t CO2)
E	industrial electricity price (USD/MWh)
LH2	levelised cost of hydrogen (USD/kg)
CF	capacity factor
Pbreak	breakeven compliance penalty (USD/t CO2)
ε	tank-to-wake emission factor (t CO2 per t MGO)
CP	power capital cost (USD/kW)
CE	energy capital cost (USD/kWh)
D	storage duration (hours)
N	annual cycle count
FOM	fixed operation and maintenance cost (USD/kW-yr)
r	discount rate
CRF	capital recovery factor.

adopted unless both failures are addressed together, a structural analogue to the multiplicative, not additive, risk modelled here. Demand-side policy is often proposed as the remedy, yet Nemet [2] finds that demand-pull incentives created by government mandate do not reliably trigger non-incremental technical change in the way market-organic demand does, because policy-created demand carries its own uncertainty about longevity. Geels' [3] multi-level account of technological transitions similarly holds that niche technologies diffuse only when several conditions align across technology, market, and institutional levels simultaneously, not when any single barrier is removed.

Functional analyses of technological innovation systems make this concrete. Bergek et al. [4] and Hekkert et al. [5] show that non-incremental technical change requires a set of system functions, including market formation, resource mobilisation and legitimation, to be simultaneously active, and that a technology can stall even when most functions are fulfilled if one remains blocked. Practitioner analyses of the hydrogen pipeline reached the same diagnosis before the cancellation wave became visible, identifying weak offtake, cost gaps and policy delay as the binding constraints and proposing demand-side remedies [51]. In the energy sector specifically, Hartley and Medlock [6] formalise this stalling as a valley of death in which technically feasible technologies fail to reach commercial viability because the price of energy from the displaced incumbent remains below the entrant's full cost even after subsidy, and Joskow [7] shows that imperfections in liberalised electricity markets can leave insufficient revenue, the missing money problem, to support capital-intensive investment even when a technology is otherwise viable.

Attempts to summarise readiness in a single metric exist, but they measure maturity rather than compound risk. The Technology Readiness Level scale tracks technical progression only. The Commercial Readiness Index (CRI) developed by the Australian Renewable Energy Agency extends the assessment beyond technical performance to a set of commercial indicators, including the regulatory environment, market opportunities, and the financial proposition [38]. It has been applied internationally as a policy design tool [39]. Both, however, are ordinal maturity ladders: a technology's position is a summary judgment on a 1-to-6 or 1-to-9 scale; the indicators are combined without an explicit aggregation rule, no uncertainty treatment is attached, and neither instrument restricts substitution between indicators so that a binding barrier can be offset by strength elsewhere, whereas the transitions literature above identifies non-compensation as the mechanism of stalling. Consequently, they cannot be used to compute the counterfactual effect of resolving one barrier versus two, which is precisely the policy question at issue. What the literature has not yet done, therefore, is score, ex ante and with an explicit, stress-testable aggregation rule, whether a technology's combination of active barriers amounts to compound failure risk, or test whether removing the most visible barrier, demand, is sufficient on its own; that is the gap this paper addresses.

This paper quantifies the risk of compound deployment. It formalises a three-tier compound-failure framework, derived from the clean-hydrogen cancellation record [8], into a reproducible Compound Deployment Risk Index (CDRI) that aggregates cause-level scores multiplicatively rather than additively, and it builds three technology-specific levelised-cost models that quantify the price gap gating demand and FID in DAC, green ammonia and LDES. Applying both to sourced data to mid-2026, it examines a proposition the framework makes explicit: that creating demand, the remedy most often proposed, is *necessary but not sufficient* to build a self-sustaining industry. The contribution is fourfold: (i) a reproducible index for compound deployment risk whose aggregation rule is stated as an explicit axiom and stress-tested across the full power-mean family of aggregators from fully compensatory to strictly conjunctive, against alternative ordinal-to-cardinal mappings of the rubric, and against coding uncertainty including correlated coder error, and which is transferable in its specification though not yet demonstrated to transfer in its calibration (Section 5.5); (ii) three cost-gap models that link the index to observable prices and penalties, providing a test of the index's structural coding that is independent of the aggregation rule; (iii) a retrodictive consistency check of the index against the archived clean-hydrogen record; and (iv) using Great Britain's LDES cap-and-floor scheme as a natural experiment, a distinction between the *sequencing* of policy, which determines whether investment happens, and its *pairing* with cost and finance, which determines whether an industry survives. Section 2 sets out the framework; Section 3 sets out the methods, data and validation; Section 4 presents the results; Section 5 discusses the design implications and limitations; Section 6 concludes.

2. The compound-failure framework

The framework is the multi-causal, three-tier failure architecture established for clean hydrogen, where more than USD 750 billion in announced investment between 2020 and 2025 yielded fewer than 5% operational projects. Systematic coding of roughly 490 confirmed cancellations to a single primary cause under a prespecified causal-primacy hierarchy shows that structural market failures dominate, carrying 74% of the cancellation population: demand absence (34%), policy and regulatory uncertainty (22%) and uneconomic production cost (18%), while systemic enabling failures (finance, permitting and the electricity cost trap) account for 22% and execution failures for 4% [8]. The architecture is summarised in Fig. 1.

Three properties distinguish this architecture from additive accounts. The cascade is *asymmetric*: Tier 1 failures trigger Tier 2 failures, not the other way around. The causes interact *multiplicatively*: because each Tier 1 cause is independently sufficient to prevent investment, capital directed at a single visible cause leaves the system intact. Moreover, the causes have different *latencies*: 1 to 3 years for rulemaking and 10 to 15 years for demand-market development. A demand-side mechanism interrupts the demand node, but because cost, financial, and technical factors persist, single-node resolution leaves the loop intact. Breaking two nodes is not claimed to be sufficient either: under the conjunctive logic, any remaining binding cause continues to gate investment, and the demand-plus-cost counterfactual of Section 4.1 accordingly still leaves DAC and green ammonia near 0.5. The claim is ordinal, that dual-node resolution moves the index materially further than single-node resolution, not that two is a sufficiency threshold. Note

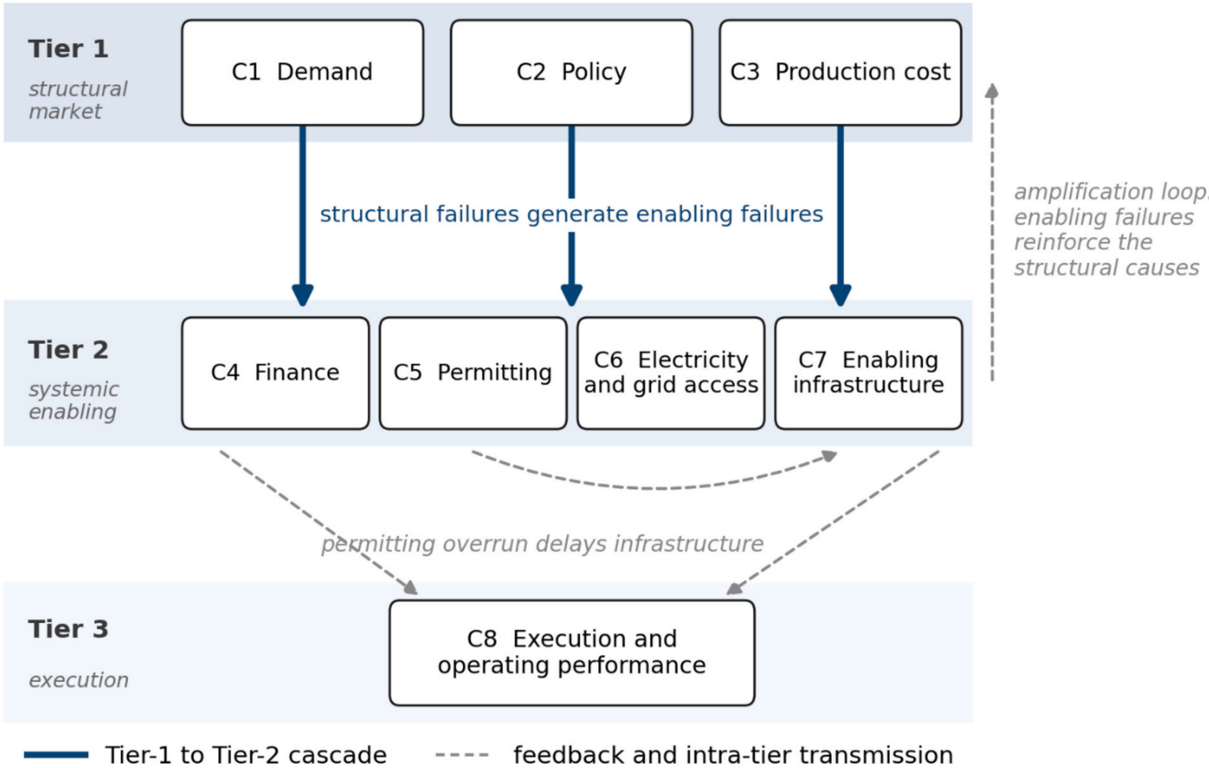

Fig. 1. Conceptual schematic of the compound failure architecture. Heavy arrows denote the Tier-1 to Tier-2 cascade; light arrows denote feedback within and between Tiers 2 and 3, including permitting, overrun, and enabling infrastructure. The tier assignment reflects the causal position rather than the cost magnitude (Section 2); the figure presents the framework, not a result.

that the asymmetric cascade implies that the causes are causally linked rather than separable, so the product form used in the index does not represent the architecture exactly; what that costs the index, and what it does not license, is set out in Section 3.1.2 (Figs. 2 and 3).

The tier assignment requires explicit criteria because three of the boundaries are not self-evident, and one appears, at first glance, to be contradictory. The criterion applied throughout is *causal position, not magnitude*: a cause is Tier 1 if it determines whether a viable transaction exists at all; Tier 2 if it determines whether an otherwise viable transaction can be financed, consented to, and physically delivered; and Tier 3 if it arises only after commissioning. Tier assignment therefore says nothing about how large a cause's cost impact is, and the following delineations follow from it.

2.1. Production cost (C3, Tier 1) against electricity cost and grid access (C6, Tier 2)

Electricity is the dominant input cost for green hydrogen and

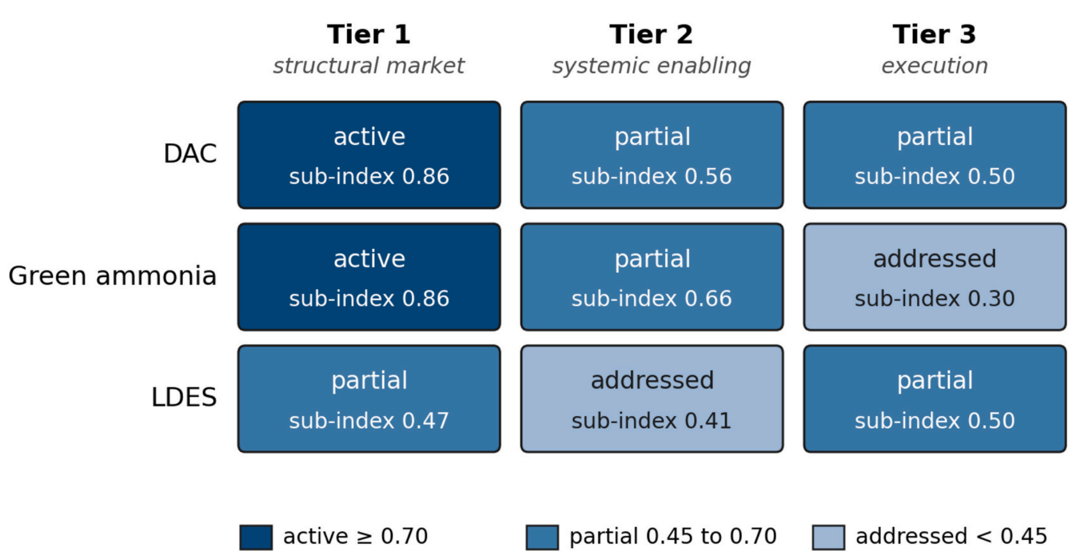

Fig. 2. Conceptual schematic of tier activation across DAC, green ammonia and LDES, 2025 to 2026. Activation states summarise the coded scores of Fig. 4 and Supplementary Table S1 and are presented for orientation rather than as independent evidence. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

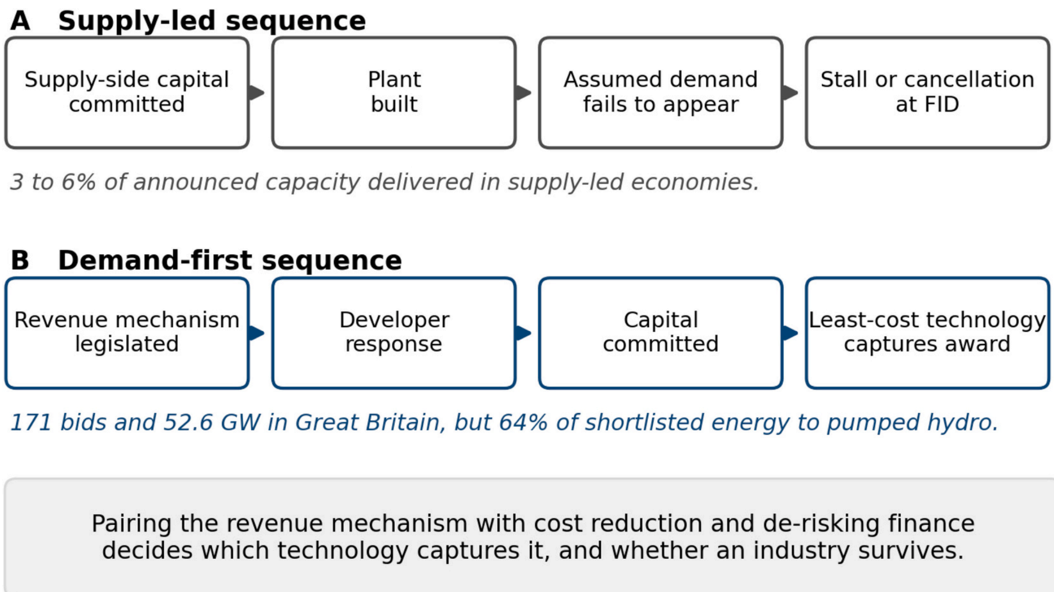

Fig. 3. Conceptual schematic of supply-led versus demand-first policy sequencing. Real jurisdictions combine both, and the two paths are drawn as limiting cases rather than as a classification (Section 5.1).

ammonia, so its placement below overall production cost looks inconsistent. It is not, under the criterion above. C3 is the aggregate delivered cost relative to the buyer's willingness to pay, and it determines whether a transaction exists; C6 is the separate question of whether the specific input can be procured at the assumed price and connected on the assumed schedule at a given site. A project may clear C3 in its business case and still fail on C6 because the connection queue exceeds the FID horizon, and the two therefore fail independently. The overlap is real, however: where a project fails because power is too expensive everywhere rather than unavailable at a site, that is a C3 failure and is coded as one. The cost models of Section 3.2 make the boundary operational, since the electricity price enters LCOD and Eq. (6) as a C3 input. In contrast, queue length and site-specific access do not enter at all.

2.2. Supply chain and enabling infrastructure (C7)

In the framework as originally derived, absent transport and storage infrastructure sat in the execution tier, which understates its effect on project economics: a DAC plant without CO2 transport and storage has no product, and this is a pre-commissioning failure of delivery rather than an operating failure. C7 is accordingly assigned to Tier 2 here, and Table 1 records that assignment. One consequence must be flagged: the tier shares quoted in the opening paragraph of this section, 74% Tier 1, 22% Tier 2 and 4% Tier 3, are computed from the archived record under the original assignment, in which infrastructure sat in Tier 3. Under the assignment used here, the Tier-2 share is therefore understated and the Tier-3 share overstated; re-derivation of the shares is listed with the recoding work in Section 5.6. The index itself is unaffected because aggregation runs across the eight causes, and the Tier-1 sub-index covers C1 to C3 under either assignment.

2.3. Execution and operating performance (C8, Tier 3) against production cost (C3, Tier 1)

These overlap whenever underperformance raises unit cost; the clearest case is a capacity factor below design, which inflates the denominator of every levelised cost. The delineation rule is temporal and counterfactual: a shortfall is coded to C8 only where the asset was built, and the shortfall emerged in operation against its own design basis; where the design basis itself was never economic at the achievable capacity factor, the failure is C3. Applied to the cases here, the Climeworks capture shortfall in Section 4.2 is C8, while the dependence of green-hydrogen cost on the capacity factor in Eq. (6) is C3.

Table 1
The eight root causes of the compound-failure framework, tier assignment and binding-level anchor conditions.

Cause	Tier	Condition defining the binding level ($a_k = 0.9$)
C1 Demand	1	No compliance obligation in force and no purchase obligation at any price inside the modelled cost range
C2 Policy and regulatory uncertainty	1	Core instrument unadopted, or enacted support subject to withdrawal or rescission within a single budget cycle
C3 Production cost	1	Modelled levelised cost exceeds every willingness-to-pay or compliance anchor across the central parameter range
C4 Finance and bankability	2	No investment-grade offtake counterparty; project WACC above 12%; no debt raised at reference projects
C5 Permitting and consenting	2	No defined consenting pathway, or consent timelines exceeding the horizon over which project economics are underwritten
C6 Electricity cost and grid access	2	Delivered power price above the level at which the cost model clears any anchor, or connection queue beyond the FID horizon
C7 Supply chain and enabling infrastructure	2	Absent transport, storage or conversion infrastructure, or sole-source supply of critical equipment
C8 Execution and operating performance	3	Demonstrated output at reference plants materially below nameplate, or repeated schedule failure

A limitation of the framework's empirical base follows directly from the coding design and should be stated before the index is built on it. The cancellation record assigns each event to a *single* primary cause under a causal-primacy hierarchy. A design of that kind can establish how often each cause appears first; it cannot establish how causes combine, because it never records two causes as jointly operative in one event. The multiplicative structure asserted here is therefore imported from the transitions literature reviewed in Section 1, where non-compensation is the theorised mechanism, and is adopted as a modelling axiom in Section 3.1.2; it is not an empirical finding read off the cancellation shares, and this paper does not claim it as one. Nor does the supporting observation that each Tier 1 cause is independently sufficient to prevent investment establish interaction, since sufficiency describes causes acting alone. Demonstrating interaction would require event-level coding, in which every operative cause is recorded, together with the joint distribution of causes across events. That recoding of the archived record is specified in the Supplementary dataset and is identified in Section 5.6 as necessary work. Until it is done, the conjunctive architecture should be read as a hypothesis that the index operationalises and that the cost models and the LDES case test indirectly, not as a result.

The framework is falsifiable, and it is useful to state in advance what evidence would count against it. Three observations would do so: (i) a technology whose coded Tier-1 causes are binding nonetheless attracting sustained, unsubsidised FIDs; (ii) a cost-gap model showing no price gap where the index codes cost as binding, or a clear gap where it codes cost as resolved; and (iii) a technology whose demand cause has been resolved by policy showing no differential deployment outcome relative to demand-unresolved peers. The cost models of Section 3.2 address (ii) directly, and the LDES case functions as the framework's positive control on (iii); the retrodictive check of Section 3.4 probes the index's calibration against the one resolved record available.

3. Methods, data and validation

3.1. Compound deployment risk index

3.1.1. Rubric and coding protocol

To make the framework operational and testable, each root cause k is assigned an activation score ak on a five-level rubric (0.1 minimal, 0.3 low, 0.5 moderate, 0.7 high, 0.9 binding). To limit coder discretion, each rubric level is behaviourally anchored to observable, documentable conditions rather than to holistic judgment. For the demand cause, for example, 0.9 (binding) is assigned only where no compliance obligation is in force and no purchase obligation exists at any price inside the modelled cost range; 0.7 where an obligation has been enacted but is not yet binding or is subject to an unresolved adoption decision; 0.5 where a binding mechanism exists but covers only part of the addressable market; and so on for each cause. The eight causes, their tier assignment and their binding-level anchor conditions are set out in Table 1. Every score in the Supplementary dataset carries its provenance: the anchor condition invoked and the primary source that documents it. This makes the coding auditable and contestable at the level of the individual score rather than of the index as a whole.

A single analyst nonetheless produced the scores, and no inter-coder reliability statistic can be reported; this is acknowledged as a limitation in Section 5.6. Two design features bound the consequences. First, the Monte Carlo treatment below perturbs every score by up to one full rubric level, which is the natural unit of disagreement between coders applying the same behavioural anchors; the reported intervals therefore already span plausible coder variation [40,42]. Second, the paper's qualitative conclusions, shown in Section 3.1.2 to be invariant to the aggregation rule and to correlation among the coding perturbations, do not hinge on fine coding distinctions. The natural extension is a structured expert judgment (SEJ) elicitation under the classical model, in which multiple calibrated experts score the rubric and are performance-weighted on calibration questions drawn from the resolved hydrogen record [40,41]; the elicitation protocol is specified in the Supplementary dataset for future application of the index.

Causes C1 to C3 constitute Tier 1 (structural market failures), C4 to C7 Tier 2 (systemic enabling failures) and C8 Tier 3 (execution failures). Only the binding level is reproduced here; the full five-level anchor set for each cause, together with the score assigned to each technology and its documentary provenance, is given in Supplementary Table S1. Activation scores by technology are plotted in Fig. 4.

3.1.2. Aggregation rule, additive alternative and dependence

Because the framework holds that all causes must be cleared for a project to proceed, and that a binding cause cannot be offset by strength elsewhere, the scores are aggregated multiplicatively rather than additively. The index is defined by Eqs. (1) to (4). Writing the pass score for cause k as.

$$\mathrm{pk} = 1 - \mathrm{ak}, \mathrm{ak} \in \{0.1, 0.3, 0.5, 0.7, 0.9\}, \mathrm{pk} \in \{0.1, 0.3, 0.5, 0.7, 0.9\} \tag{1}$$

compound viability is the product of the cause-level pass scores,

$$\mathrm{V} = \Pi\ (\mathrm{k} = 1 \text{ to } \mathrm{n})\ \mathrm{pk}, \mathrm{V} \in (0, 1] \tag{2}$$

the readiness score is its geometric mean over the n = 8 causes,

$$\mathrm{g} = \mathrm{V}\hat{}(1/\mathrm{n}), \mathrm{g} \in (0, 1] \tag{3}$$

and the Compound Deployment Risk Index is its complement,

$$\mathrm{CDRI} = 1 - \mathrm{g}, \mathrm{CDRI} \in [0, 1) \tag{4}$$

A Tier-1 sub-index is computed identically over the three structural causes ($n = 3$). The geometric mean in Eq. (3) is a scale normalisation: it returns the index to the unit interval and keeps values comparable across technologies and across sub-indices of different dimension, which the

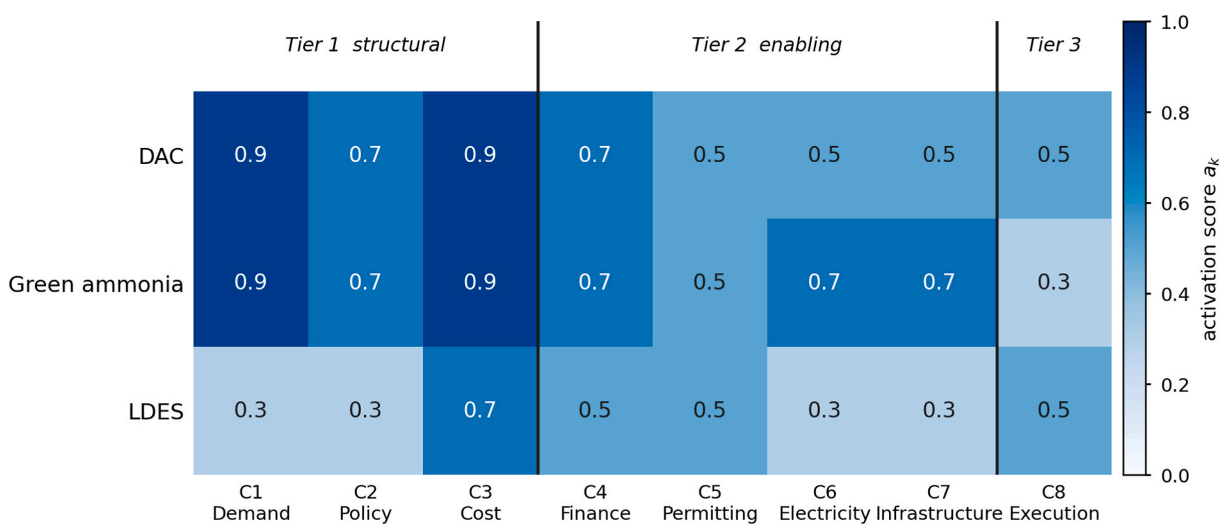

Fig. 4. Activation score ak for each of the eight causes of Table 1; the red line separates Tier 1 (C1 to C3) from Tiers 2 and 3. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

raw product V does not.

Two clarifications are required regarding interpretation, because the notation is probabilistic but the underlying measurements are not. First, ak is an ordinal rubric level, behaviourally anchored to observable conditions (Table 1), and the mapping in Eq. (1) onto the unit interval is an ordinal-to-cardinal transformation adopted for tractability, not an elicited or frequency-calibrated probability. Accordingly, pk should be read as a *pass score* and V as a *compound viability score*; neither is a probability, and the CDRI is not the probability that a project fails to reach FID. The index is an ordering instrument, and every claim made from it in this paper is comparative, either across technologies or across counterfactual states of the same technology. Second, and consequently, the numerical level of the index has no standalone meaning: a CDRI of 0.71 is interpretable only against the 0.44 of LDES and the 0.71 of clean hydrogen at its announcement peak (Section 3.4), not against an absolute failure threshold. The sensitivity of the index to the choice of mapping is reported in Section 3.1.3, and calibration onto an elicited probability scale via structured expert judgment is set out as the natural extension in Section 3.1.1.

Multiplicative aggregation is a modelling axiom, not an empirical finding: it imports the framework's non-compensatory logic directly into the index, thereby distinguishing the CDRI from additive composite indicators [42] and from maturity ladders such as the TRL and the CRI [38,39]. Because the paper's headline counterfactual behaviour is partly a consequence of this axiom (Section 3.1.4), Section 3.1.3 justifies the choice theoretically and stress-tests it by examining what the causal linkage among the causes does and does not imply for the index, as treated below.

3.2. Causal linkage among the causes

The framework holds that Tier-1 failures generate Tier-2 failures, so the eight causes are causally linked rather than separable, and the product in Eq. (2) does not represent that linkage. An earlier treatment of this point invoked Fréchet-Hoeffding bounds on non-negatively associated pass events and concluded that the reported CDRI is an upper bound on compound risk. That argument is withdrawn. Copula bounds are results about joint probability distributions [43], whereas pk is a pass score derived from an ordinal rubric and not a calibrated probability, and V is correspondingly not a joint probability. Applying probabilistic dependence results to these quantities would import precisely the probabilistic reading that the preceding paragraph disclaims, and no bound on compound risk in the probabilistic sense can be read off the index.

What survives is an algebraic sensitivity rather than a probabilistic bound. Replacing the product in Eq. (2) with V = min over k of pk, and normalising as in Eq. (3), gives an index that responds only to the worst cause and therefore represents the limiting case in which the causes overlap completely in what they measure; it returns 0.25, 0.25 and 0.14 for DAC, green ammonia and LDES, preserving the ranking (Supplementary Table S4). It is reported for that reason alone, and it is not independent corroboration: like the $s \to -\infty$ column of Table 2, it is a monotone function of min over k of pk, so the two are the same sensitivity expressed under different normalisations, and neither carries information beyond the worst cause. The direction of the substantive concern is nonetheless worth stating plainly. If causal linkage means that the coded causes partly measure the same underlying condition, then the product counts that condition more than once, and the index level is overstated. This is a further reason, alongside the ordinal-to-cardinal mapping discussed above, why no absolute meaning attaches to the level, and why every claim made from the index in this paper is comparative.

A separate, narrower exercise addresses correlation within the *coding* rather than among deployment outcomes. The Monte Carlo of Section 3.1.4 perturbs each score independently, which assumes that a coder's errors are unrelated across causes, whereas a single analyst applying one rubric is more likely to err in a common direction. A Gaussian copula with exchangeable pairwise rank correlation ρ in {0, 0.3, 0.6} is therefore imposed on the perturbations. The cross-technology ordering (DAC and green ammonia above LDES) and the ordering of the two counterfactuals are preserved at every tested ρ: medians are essentially unchanged (0.69, 0.72 and 0.45 to 0.46) while the 90% intervals widen to [0.53, 0.80], [0.56, 0.83] and [0.29, 0.61] at ρ = 0.6 (Supplementary Table S5). This establishes that the reported intervals are robust to correlated coding error. It says nothing about dependence among real deployment events, which the single-primary-cause design of the underlying record cannot identify (Section 2) and which the recoding described in Section 5.6 would be required to estimate.

Table 2
Compound Deployment Risk Index under the power-mean family of aggregators.

Aggregator	s	Substitution between causes	DAC	Green ammonia	LDES
Arithmetic mean (additive)	1	Full (compensatory)	0.65	0.68	0.43
Power mean	0.5	High	0.68	0.70	0.44
Geometric mean (this paper)	**0 (limit)**	**Unit elasticity**	**0.71**	**0.73**	**0.45**
Power mean	−1 (harmonic)	Low	0.77	0.78	0.47
Power mean	−5	Very low	0.87	0.87	0.57
Minimum (strictly conjunctive)	−∞	None	0.90	0.90	0.70

3.2.1. Justification and stress-testing of the aggregation rule

The choice of multiplicative aggregation requires justification against other nonlinear rules, and not merely against an additive alternative. The justification offered here is that the choice is not between the geometric mean and an arbitrary alternative, but between degrees of *compensation*, and that the paper's conclusions do not depend on where in that range the aggregator sits. Both points are made by embedding Eq. (3) in the one-parameter power-mean family.

$$gs = [(1/n)\ \Sigma\ (k = 1 \text{ to } n)\ pk\hat{}s]\hat{}(1/s), s \in [-\infty, 1] \tag{5}$$

whose parameter s is the elasticity of substitution between causes [44]. At s = 1, the rule is the arithmetic mean, and causes are fully compensatory: a cleared permitting regime can offset an absent market. As s falls, compensation declines; the limit $s \to 0$ recovers the geometric mean of Eq. (3), and the limit $s \to -\infty$ recovers the minimum, under which a technology is exactly as ready as its worst cause and no compensation is possible at all. The transitions literature reviewed in Section 1 [3–5] asserts non-compensation, that a single blocked system function stalls diffusion irrespective of the others, and therefore locates the admissible aggregator in the range $s \le 0$. It does not identify a point in that range. The geometric mean is selected within it on three grounds: it is the unique power mean that is invariant to the ordinal-to-cardinal rescaling discussed in Section 3.1.2 up to a monotone transformation, it remains sensitive to every cause rather than to the worst alone, and it is the standard non-compensatory rule in the composite-indicator literature [42].

Table 2 reports the index under the whole family. The cross-technology ordering, DAC and green ammonia materially above LDES, is invariant at every value of s from fully compensatory to strictly conjunctive, as is the ordering of the two counterfactuals. What the choice of s changes is the *level* of the index and the size of the counterfactual response, not any comparative conclusion drawn in this paper. A reader who rejects the axiom entirely may read the s = 1 column, which is the additive counterpart CDRIadd = (1/n) Σ ak reported throughout the Supplementary dataset (Table S3; 0.65, 0.68 and 0.43 at baseline, falling to 0.55/0.58/0.40 under demand-only and 0.45/0.48/0.33 under demand-plus-cost resolution), and still recover the paper's comparative conclusions.

The parameter s in Eq. (5) governs substitutability between causes: $s = 1$ is fully compensatory, $s \to 0$ is the geometric mean used throughout the paper, and $s \to -\infty$ is the strictly conjunctive minimum. The cross-technology ordering is preserved at every value. Counterfactual values under each rule are given in Supplementary Table S3.

Two further properties should be stated plainly, because they qualify how strongly the conjunctive claim can be made. First, the geometric mean is *non-compensatory but not strictly conjunctive*. A technology with one binding cause and seven cleared causes scores 0.32 under Eq. (3), which is below the LDES baseline of 0.44, whereas under the minimum rule it scores 0.90. The index therefore does not implement the literal proposition that any single binding barrier is sufficient to prevent investment; it implements the weaker and, it is argued here, more defensible proposition that binding barriers are only partially substitutable and that their effects compound. Readers who wish to impose the literal proposition should use the $s \to -\infty$ column of Table 2, which preserves every ranking reported here. The earlier framing of the CDRI as encoding strict conjunction is accordingly withdrawn in favour of non-compensation.

Second, the index inherits whatever cardinal content the rubric mapping supplies. *Re*-running the index under alternative mappings of the five rubric levels onto the unit interval, a wider map {0.05, 0.275, 0.5, 0.725, 0.95} and a narrower map {0.2, 0.35, 0.5, 0.65, 0.8}, moves the DAC index between 0.64 and 0.76 and the green-ammonia index between 0.66 and 0.78. At the same time, LDES stays between 0.44 and 0.46. The separation between the two high-risk technologies and LDES holds across all tested mappings, but the levels do not, which is a further reason to treat the index ordinally (Section 3.1.2). Full results are in Supplementary Table S3b.

Finally, the CDRI can be compared with established multi-criteria decision-making methods applied to the same eight-cause matrix. A TOPSIS ranking [45] and an equal-weight additive value function of the kind underlying AHP [46] both reproduce the cross-technology ordering (LDES least risky, DAC intermediate, green ammonia most). This is a useful convergent check, and it also locates the CDRI's contribution precisely: TOPSIS and AHP are compensatory by construction, sitting at $s = 1$ in Eq. (5), and they return a ranking rather than a structured object, so neither supports the node-resolution counterfactuals of Section 3.1.4 nor distinguishes a technology that clears every tier from one that clears only the demand tier. The CDRI is best understood not as a competitor to these methods but as a domain-specific aggregation restricted to the non-compensatory range and paired with an explicit causal architecture.

3.2.2. Counterfactuals: what they quantify and what they do not test

Two counterfactuals are computed for each technology: resolving demand alone (setting the demand score to 0.1) and resolving demand and cost together (setting both to 0.1). Their interpretation requires care. Given the multiplicative axiom, the qualitative pattern that resolving one of eight causes moves the index less than resolving two follows arithmetically from the aggregation rule; the counterfactuals are therefore not an independent statistical test of the necessary-but-not-sufficient proposition but a quantification of its magnitude given the observed score profile. Their informative content lies in the sizes: how far single-node and dual-node resolution move each technology, given where its remaining causes actually sit.

The empirical tests of the proposition lie elsewhere, in four places. First, the coded activation scores themselves, which are anchored to observable market and regulatory facts and are individually falsifiable. Second, the cost-gap models of Section 3.2, which test whether the price gap the index codes as binding is in fact observable in market prices and penalties, independently of any aggregation rule. Third, the LDES natural experiment of Section 4.4, in which the demand cause was resolved by policy while cost and finance were not: an out-of-index observation of what single-node resolution produces in the field. Fourth, the retrodictive check of Section 3.4 against the resolved clean-hydrogen record.

Uncertainty in the coding is propagated by Monte Carlo simulation. Each score is perturbed by up to one rubric level (uniform plus or minus 0.2, clipped to [0.1, 0.9]) over 10,000 draws, and 5th to 95th percentile intervals are reported (random seed 42).

3.3. Cost-gap models

Three reduced-form levelised-cost models, each analogous to the levelised-cost-of-hydrogen (LCOH) model in the companion analysis [8], quantify the price gap that gates demand and FID. All use the capital recovery factor CRF(r,n) = r(1 + r)^n / [(1 + r)^n − 1]. The models are deliberately reduced-form: their purpose is to establish whether an order-of-magnitude price gap exists at the point where a buyer or a compliance authority decides, not to reproduce a process-engineering cost estimate for any particular plant. Where that simplification is material, it is stated below, and Section 3.5 propagates it by jointly varying all parameters. Central values, ranges, and parameter-level sources are provided in Table 3 and the Supplementary dataset.

(i) *DAC*. The levelised cost of removal is $\text{LCOD} = A + \eta E \cdot E$, where A is the annualised non-energy cost per tonne of CO_2 removed, and ηE is the electricity intensity in MWh per tonne. The system boundary is capture, regeneration, on-site conditioning and compression to pipeline specification. A aggregates capital recovery at a 12% weighted average cost of capital over a 20-year life, fixed operation and maintenance, and sorbent or solvent replacement; central 400 USD/t, band 200 to 850 USD/t. No DAC-specific study of project financing costs is available. The 12% figure is adopted as a first-of-a-kind climate-infrastructure benchmark by analogy with the financing-cost evidence assembled for green hydrogen in emerging and developing economies [20], which concerns a different technology and is used here only as an order-of-magnitude anchor; the discount rate is varied over 8 to 15% in the Supplementary dataset, and the LCOD band reported below spans that variation. The band deliberately spans nth-of-a-kind mature plants at the lower bound and first-of-a-kind plants at the upper bound, bracketing the published techno-economic range for both principal process families, low-temperature solid sorbent and high-temperature liquid solvent [10,47,48]; the central value corresponds to a mature commercial-scale plant at an 85% capacity factor. Because A is the dominant term and sets both the position and the width of the band in Fig. 6, it is varied over its full range in the joint uncertainty analysis of Section 3.5.

$\eta E = 2.0$ MWh/t central, range 1.5 to 2.5, is a total energy intensity expressed in electricity-equivalent terms. This is the model's principal simplification, adopted for a specific reason. The two process families differ mainly in the *form* of energy they require: solid-sorbent systems are driven largely by low-grade heat at 80 to 120 °C plus fans and vacuum. In contrast, liquid-solvent systems require high-temperature calcination at 900 °C or higher [47,48]. Separately resolving thermal demand, heat integration, and process configuration would require a configuration-specific model and a site-specific heat price. They would not change the finding at issue, because the gap between LCOD and the compliance anchors is dominated by A rather than by the energy term: at the central A, the entire energy contribution across the full 1.5 to 2.5 MWh/t and 40 to 120 USD/MWh ranges spans 60 to 300 USD/t, against an anchor set topping out at 180 USD/t; setting the energy term to zero altogether would still leave the central A above every compliance anchor. Converting thermal demand to electricity-equivalent is nonetheless conservative in one direction and not in the other: where genuine waste heat is available at near-zero cost, the model overstates cost for solid-sorbent plants; where heat is electrified or purchased, it is accurate. A configuration-resolved DAC cost model, distinguishing sorbent and solvent pathways with separate thermal and electrical intensities and an explicit heat price, is identified in Section 5.6 as a priority extension.

Three exclusions should be stated explicitly. CO_2 transport,

Table 3
Cost-model parameters, central values, ranges and parameter-level sources.

Model	Parameter	Central	Range	Source
DAC	A, annualised non-energy cost (USD/t CO_2)	400	200–850	[10,20,47,48]
DAC	ηE, energy intensity, electricity-equivalent (MWh/t CO_2)	2.0	1.5–2.5	[9,10,47,48]
DAC	E, industrial electricity price (USD/MWh)	80	40–120	[11]
DAC	WACC / economic life	12% / 20 yr	–	[20]
Green ammonia	Electrolyser electricity use (kWh/kg H_2)	55	–	[8]
Green ammonia	CF, capacity factor (fraction)	0.50	0.40–0.65	[8]
Green ammonia	Conversion cost, Haber-Bosch + ASU (USD/t NH_3)	180	100–300	[14]
Green ammonia	H_2 requirement (kg per t NH_3)	178	–	[14]
Green ammonia	LHV NH_3 / MGO (GJ/t)	18.6 / 42.7	–	[14]
Green ammonia	MGO price (USD/t)	650	550–780	[14]
Green ammonia	ε, tank-to-wake emission factor (t CO_2/t MGO)	3.2	–	[14,49]
LDES	CP, power capital cost (USD/kW)	250 Li-ion; 600 VRFB; 800 Fe-air; 1000 CAES; 1400 PSH	±30%	[15,16,50]
LDES	CE, energy capital cost (USD/kWh)	200 Li-ion; 180 VRFB; 20 Fe-air; 50 CAES; 25 PSH	±30%	[15,16,50]
LDES	RTE, round-trip efficiency	0.86 Li-ion; 0.70 VRFB; 0.50 Fe-air; 0.55 CAES; 0.80 PSH	–	[15,16,50]
LDES	N, annual cycles	365 Li-ion, VRFB; 200 PSH, CAES; 100 Fe-air	equal-N case reported	[15,16,50]
LDES	Pchg, charging price (USD/MWh)	40 common	20–60; differentiated case in Section 4.4	[15,16]
LDES	r, discount rate	8%	–	[15,16,50]

geological storage and monitoring, reporting and verification are outside the boundary and would add roughly 15 to 40 USD/t; lifecycle emissions from construction and energy supply, which reduce net removal per tonne captured, are not deducted; and no allowance is made for capacity-factor shortfall of the kind documented at operating plants in Section 4.2. All three omissions widen the gap the model reports, so the LCOD estimates here are a lower bound on the true cost of a delivered, verified net removal. Consistent with that boundary, LCOD is a cost per tonne *permanently stored*; where the text refers to gross capture, to contracted volumes or to net removal, the distinction is stated. LCOD is compared with willingness-to-pay anchors: the EU Emissions Trading System allowance (65–90), the 45Q credit for DAC with permanent geological storage (180), and voluntary durable-removal prices (300–600 USD/t) [11,13].

(ii) *Green ammonia.* The delivered cost is built from the green-hydrogen levelised cost of the companion analysis [8],

$$LH2 = 0.055 \cdot E + 0.481/CF + 0.45 \ (\text{USD/kg } H_2) \tag{6}$$

where E is the industrial electricity price in USD/MWh, and CF is the capacity factor entered *as a fraction* on (0, 1]. The coefficient is stated here as 0.481 rather than the 48.1 in [8], which reported CF in percentage points; the two forms are algebraically identical, and the change is made solely to remove an ambiguity in units. The three terms are the electricity cost at 55 kWh per kg, the annualised electrolyser capital charge, and variable operation and maintenance. Ammonia is then costed at 178 kg H_2 per tonne NH_3 with a Haber-Bosch and air-separation conversion cost of 180 USD/t (range 100 to 300), and expressed on a marine-gas-oil-equivalent (MGOe) basis using lower heating values of 18.6 and 42.7 GJ/t [14]. The green premium over conventional fuel (650 USD/t MGO) is converted to a breakeven compliance penalty Pbreak = premium / ε, with ε = 3.2 t CO_2 per tonne MGO; fuel-switching is economic only where the applied penalty exceeds Pbreak. Three scenarios are reported: favourable (E = 40 USD/MWh, CF = 0.60), central (E = 60 USD/MWh, CF = 0.50), and European (E = 80 USD/MWh, CF = 0.45).

The conversion cost is held fixed across scenarios, which suppresses its dependence on plant scale, synthesis pressure, air-separation technology and, most importantly, on the capacity factor of the synthesis loop, since Haber-Bosch is capital-intensive and disfavours intermittent operation. The 100 to 300 USD/t range is intended to bracket this variation and is sampled independently in Section 3.5; there it contributes far less to the spread of Pbreak than the electricity price and capacity factor, because a 200 USD/t swing in conversion cost moves the breakeven penalty by about 143 USD/t CO_2. In contrast, the E and CF ranges together move it by over 300.

Two boundary caveats attach to the comparison with the IMO mechanism, and both are reasons to read Pbreak as an indicative screening estimate rather than a model of a shipowner's compliance decision. First, ε is a tank-to-wake factor, whereas the Net-Zero Framework regulates well-to-wake lifecycle GHG intensity under separate guidelines [49]; the two boundaries are not equivalent, and the effective breakeven for any particular cargo depends on the certified upstream intensity of the specific ammonia pathway. Second, the framework is not a binary choice between paying a remedial-unit price and buying green ammonia: surplus units, compliance pooling and balancing across a fleet, and the reward tier for over-compliance all enter, as do propulsion-efficiency differences, pilot-fuel requirements, ammonia slip, N_2O formation and onboard storage volume, none of which the LHV-equivalence treatment captures. These omissions do not all point the same way. However, the largest of them, the N_2O and slip penalties on the lifecycle intensity of ammonia combustion, raise rather than lower the effective breakeven, so the direction of the reported gap is robust even though its magnitude is approximate.

(iii) *LDES.* The levelised cost of storage is.

$$LCOS = 1000 \cdot [CRF \cdot (CP + D \cdot CE) + FOM]/(D \cdot N) + Pchg/RTE \ (\text{USD/MWh}) \tag{7}$$

where CP and CE are power (USD/kW) and energy (USD/kWh) capital costs, D is duration in hours, N is the annual cycle count, FOM is fixed O&M in USD/kW-yr, Pchg is the charging electricity price, and RTE is round-trip efficiency; the factor of 1000 converts the per-kW and per-kWh terms to USD/MWh. Technology parameters are taken from public cost-and-performance assessments [15,16,50] and reported in Table 3.

Two modelling choices in the original specification required correction and are treated explicitly here. First, the frontier long-duration category is no longer represented by a single composite chemistry. Iron-air, vanadium redox flow and compressed air differ so sharply in round-trip efficiency (0.50, 0.70 and 0.55 respectively) and in the ratio of power to energy capital cost that a single representative technology is not defensible; each is therefore modelled separately alongside lithium-ion and pumped storage hydro, and Fig. 8 reports all five. Second, the annual cycle count is technology-specific rather than common, since the economic case for long-duration assets rests on infrequent deep cycling, and imposing a common N would flatter the incumbent; N is set at 365 for lithium-ion and flow, 200 for pumped hydro and compressed air, and 100 for iron-air, and an equal-cycle counterfactual is reported in the Supplementary dataset.

The charging electricity price is the most consequential remaining assumption. Because Pchg enters Eq. (7) divided by RTE, a common charging price systematically favours the high-efficiency incumbent, and in real system operation charging opportunities are not common across technologies: an asset cycling a hundred times a year can select the cheapest hours of the year, whereas a daily-cycling asset cannot. Two cases are therefore reported. In the central case, Pchg = 40 USD/MWh for all technologies, with sensitivity across 20–60 USD/MWh (Supplementary Fig. S2). In a technology-differentiated case, Pchg is set to reflect the surplus-energy hours each technology can realistically access (15 USD/MWh for iron-air, 20 USD/MWh for compressed air, 25 USD/MWh for pumped hydro, 30 USD/MWh for flow, 40 USD/MWh for lithium-ion). Section 4.4 reports both, because the two cases do not give the same answer, and the difference between them is itself a policy finding.

Every parameter used in Figs. 6–8 is listed with the reference supporting that specific value; composite citations at the end of a paragraph are no longer used for parameter attribution. Ranges are the support of the distributions sampled in the joint uncertainty analysis of Section 3.5. Full derivations, the distributional forms assumed, and the analysis script are in the Supplementary dataset.

3.4. Data

The analysis uses agency and regulatory data (IEA, IMO, Ofgem, U.S. DOE), techno-economic parameters from peer-reviewed and national-laboratory assessments, and primary company disclosures, for the period to mid-2026; trade-press items are used only to corroborate primary company or regulatory sources, and no score or model parameter rests on a trade-press item alone. All parameters, rubric scores with provenance, and computed tables are deposited (see Data availability).

3.5. Retrodictive consistency check against the clean-hydrogen record

The framework was derived from the hydrogen cancellation record, so hydrogen cannot serve as a fully independent test of the index; it can, however, serve as a retrodictive consistency check of the index's calibration, because the rubric scores used here were not fitted to that record. Clean hydrogen is coded at two vintages using the same behaviourally anchored rubric: 2021, near the peak of announcement activity, using the sources available at that date, and 2025, after the cancellation wave documented in [8]. The check asks two questions. First, does the 2021-vintage index place clean hydrogen in the same high-risk band that the realised record implies ex post? The 2021-vintage CDRI is 0.71, against a 2025-vintage value of 0.76 (coding and provenance in Supplementary Table S6): the index places clean hydrogen at its announcement peak squarely in the band the realised record implies ex post, identical to DAC (0.71) and adjacent to green ammonia (0.73) today. Second, does the index ordering across causes align with the realised cause-level cancellation shares in the archived dataset? Over the four causes with published shares (demand 34%, policy 22%, cost 18%, execution 4%), the Spearman rank correlation is 0.63; with four observations, this carries descriptive rather than inferential weight, and the full eight-cause test, requiring the Tier-2 share split, is specified in the Supplementary dataset. The exercise is therefore a consistency check on the index's calibration, as described throughout; it is not a validation. Two reasons preclude the stronger word. The framework was derived from the same cancellation record, so the test is not independent. And although the 2021 vintage was coded only from sources available at that date, it was coded in 2026 by an analyst who knew the outcome, so hindsight bias cannot be excluded by design. A genuinely ex-ante test, scoring a cohort of technologies today and evaluating realised FID outcomes at a fixed horizon, is beyond the scope of a single paper and is identified in Section 5.6 as the decisive next step for the instrument.

3.6. Joint techno-economic uncertainty

The counterfactual analysis in Section 3.1.4 is deterministic with respect to the cost parameters, whereas the scenarios in Section 3.2 vary parameters one at a time. Neither establishes whether the reported gaps survive simultaneous variation. A joint Monte Carlo is therefore run over all cost-model parameters together (200,000 draws, seed 42): for DAC, A as a triangular distribution on (200, 400, 850), ηE uniform on (1.5, 2.5) and E uniform on (40, 120); for green ammonia, E uniform on (40, 90), CF uniform on (0.40, 0.65), conversion cost triangular on (100, 180, 300), the hydrogen requirement normal about 178 kg, the MGO price uniform on (550, 780) and ε normal about 3.2. Results are reported as the probability that each cost gap closes, which is the decision-relevant quantity, and are given in full in Supplementary Table S8. Because the draws are over parameter uncertainty rather than over realised market outcomes, these probabilities describe the model's sensitivity, not a forecast.

4. Results

4.1. Compound deployment risk index across the three technologies

The index separates the three technologies clearly (Figs. 4 and 5, Table 4). DAC and green ammonia score 0.71 and 0.73 (Monte-Carlo medians 0.69 and 0.72; 90% intervals [0.63, 0.75] and [0.66, 0.78]), while LDES scores 0.44 (median 0.46, [0.38, 0.54]). The gap is driven by Tier 1: the structural sub-index is 0.86 for both DAC and green ammonia, compared with 0.47 for LDES, reflecting that only LDES has a binding demand mechanism. The separation is robust to every treatment of Sections 3.1.2 and 3.1.3: it holds at every value of the aggregation parameter s from fully compensatory to strictly conjunctive (Table 2), under all four rubric mappings tested (Supplementary Table S3b), under the minimum-based variant of Eq. (2) (0.25, 0.25 and 0.14; Supplementary Table S4), under correlated coding-error perturbations at all tested values of ρ with essentially unchanged medians (Supplementary Table S5), and under compensatory MCDM rules applied to the same score matrix (Section 3.1.3).

The counterfactuals quantify the framework's central implication given the observed scores (Fig. 5, Table 4). Resolving the demand cause alone lowers the DAC index by 0.09 (from 0.71 to 0.61) and the green-ammonia index by 0.09 (from 0.73 to 0.64), because cost and the other structural causes remain binding and the multiplicative structure keeps viability low. Simultaneously resolving demand and cost lowers them by 0.22 and 0.20 (to 0.49 and 0.53), approaching the LDES range. As Section 3.1.4 sets out, the direction of this pattern is an arithmetic consequence of the aggregation axiom. It is not evidence for it: changing two inputs necessarily moves the index further than changing one. What the counterfactual supplies is the magnitude: under 0.10 for single-node resolution, compared with about 0.20 for dual-node resolution, given where the remaining coded causes actually sit; and even after dual-node resolution, both technologies remain near 0.5, so 2 is not a sufficiency threshold. Whether single-node resolution is in fact insufficient in the

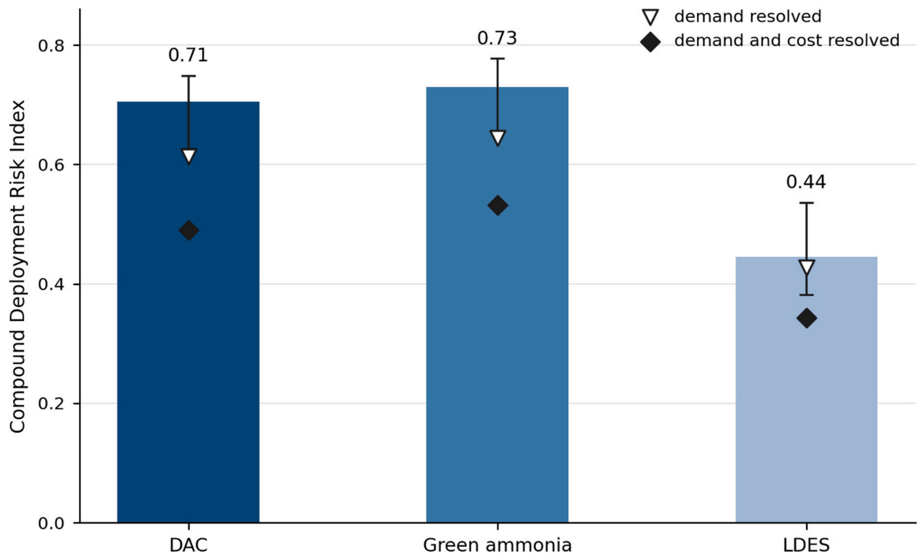

Fig. 5. CDRI by technology with 90% Monte-Carlo intervals (bars), and the counterfactuals resolving demand alone (down-triangles) and demand with cost (diamonds).

Table 4
CDRI baseline, Tier-1 sub-index, Monte-Carlo interval and demand-intervention counterfactuals.

Technology	CDRI	Tier-1 CDRI	MC median [5%, 95%]	Demand only	Demand + cost
DAC	0.71	0.86	0.69 [0.63, 0.75]	0.61	0.49
Green ammonia	0.73	0.86	0.72 [0.66, 0.78]	0.64	0.53
LDES	0.44	0.47	0.46 [0.38, 0.54]	0.43	0.34

field is a separate question, tested independently by the LDES natural experiment (Section 4.4), the cost-gap models (Sections 4.2 and 4.3) and the retrodictive check (Section 3.4). Throughout what follows, the counterfactual magnitudes are reported as properties of the index and the field evidence as the test of the proposition; the two are not merged.

Counterfactual columns show the CDRI after resolving the demand cause alone and after resolving demand and cost together, at the minimal rubric level. The additive-index counterparts of every column, and the minimum-based variant of Eq. (2), are given in Supplementary Tables S3 and S4. Full scores and computations are in the Supplementary dataset.

4.2. Direct air capture: the demand gap

Structural demand absence is confirmed quantitatively by the removal-cost gap (Fig. 6). Central LCOD ranges from 480 to 640 USD/t across industrial electricity prices of 40 to 120 USD/MWh and reaches 930 to 1090 USD/t for first-of-a-kind plants; the electricity component alone contributes 160 USD/t at 80 USD/MWh. The comparison with willingness-to-pay anchors must be stated in terms of the anchor rather than in aggregate. Against the compliance anchors, the gap does not close anywhere in the parameter space: the minimum modelled LCOD, at the joint lower corner of Table 3 (A = 200 USD/t, ηE = 1.5 MWh/t, E = 40 USD/MWh), is 260 USD/t, and in the joint uncertainty analysis of Section 3.5 the probability that LCOD falls below either anchor is zero across 200,000 draws. The two comparisons are not equally strong, however, and an earlier undifferentiated statement that no parameter combination brings LCOD within a factor of two of either anchor was incorrect. Relative to the 45Q credit the shortfall is a factor of 1.4 at that lower corner and 2.7 to 3.6 at the central A, so the credit does sit within a factor of two of modelled cost in one corner of the parameter space, although only 0.9% of joint draws fall below 360 USD/t. Relative to the EU ETS allowance, the shortfall is 2.9 at the lower corner and 5.3 to 9.8 at the central A, and this comparison alone approaches an order of magnitude. What holds without qualification is the narrower statement, which is the one the argument requires: no parameter combination brings modelled LCOD below either compliance anchor. Against voluntary durable-removal prices, the picture is different, and the earlier formulation was too strong. The central LCOD line crosses the top of the modelled voluntary anchor band (600 USD/t) at an industrial electricity price of 100 USD/MWh, so at 40 to 100 USD/MWh a mature central-case plant sits inside, not above, the upper voluntary band; across the joint draws LCOD is below 600 USD/t in 42% of cases and below the mid-range 450 USD/t in 8%. The DAC demand problem at the top of the voluntary market is therefore a *volume* constraint rather than a price constraint, which is precisely the thinness documented below: the segment that can pay 600 USD/t exists but is small and buyer-concentrated. It is the absence of any *compliance* anchor at a price the technology can reach, combined with that thinness, that leaves the demand cause binding. No jurisdiction has a mandatory purchase obligation to close the gap.

One qualification to the demand coding deserves separate treatment, because it is the strongest existing counter-case and it runs against the argument made here. DAC demand need not be standalone. Integrated facilities pairing DAC with electrolytic hydrogen produce electrofuels, and for aviation the European Union has legislated binding demand for exactly that product: ReFuelEU Aviation imposes a synthetic-fuel sub-mandate beginning at 1.2% of aviation fuel supplied in 2030 and rising through the 2040s, with delegated rules progressively restricting the eligibility of CO2 captured from fossil point sources so that atmospheric and biogenic CO2 become the compliant feedstocks [52]. That is a compliance obligation, enacted and in force, whose derived demand reaches DAC through the electrofuel chain; and vertical integration does comparable work for ammonia, where a producer holding both the electrolysis and synthesis steps, or combining fertiliser and marine off-take, is not exposed to the merchant green premium in the way the standalone breakeven of Section 4.3 implies. The NEOM arrangement in Section 4.3, where a single offtaker takes up to 1.2 Mtpa and a second party commercialises the remainder, is a partial instance of this.

Two things follow. First, the coding is affected. Under the rubric anchors of Table 1, an obligation that is enacted but not yet binding, or that covers only part of the addressable market, is scored 0.7 rather than

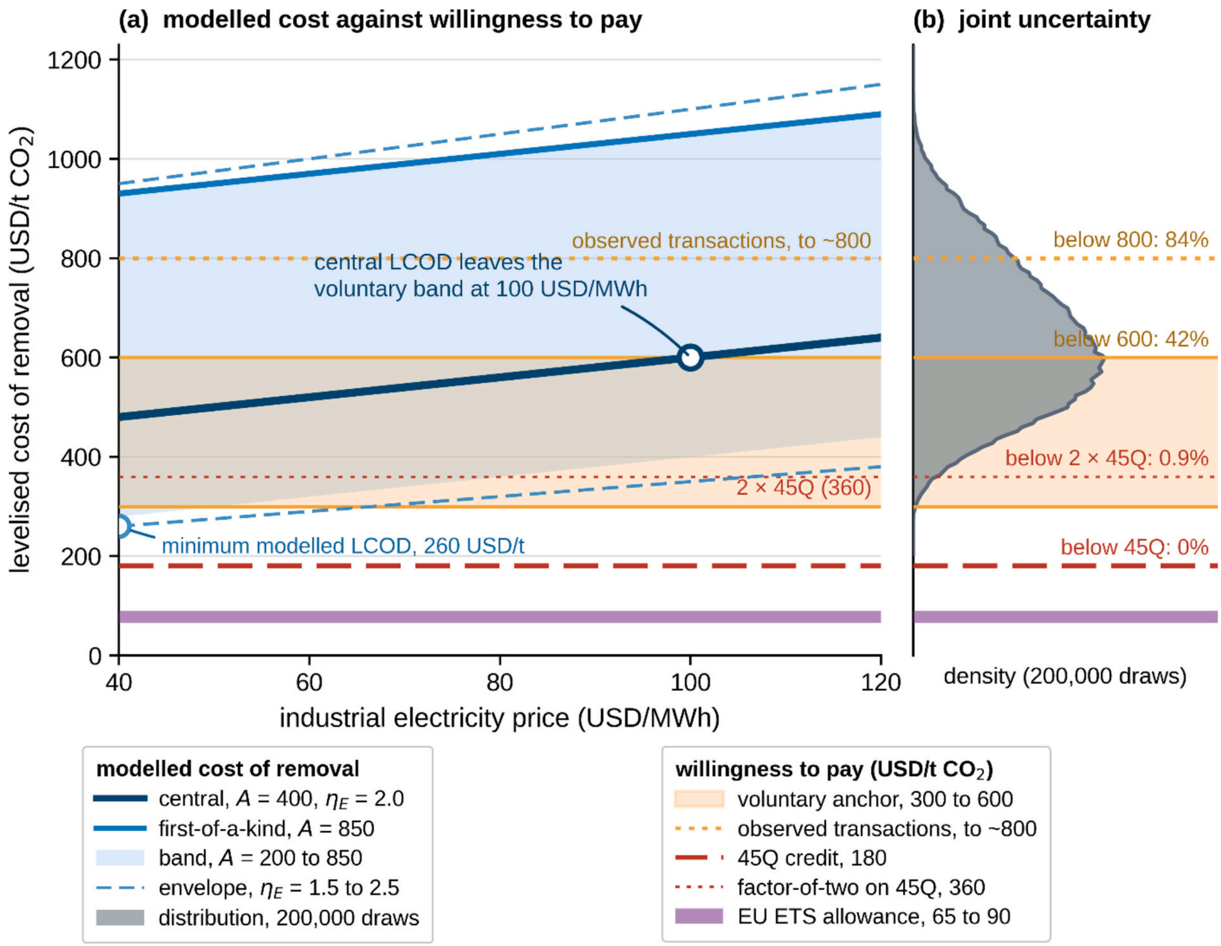

Fig. 6. DAC levelised cost of removal against willingness-to-pay anchors. (a) LCOD as a function of the industrial electricity price. The heavy dark line is the central case (A = 400 USD/t, ηE = 2.0 MWh/t) and the lighter solid line the first-of-a-kind case (A = 850 USD/t); the pale blue band spans the modelled range of A at ηE = 2.0, and the dashed lines give the full envelope obtained by varying ηE over 1.5 to 2.5 as well (Table 3). The open marker at the left edge is the minimum modelled LCOD, 260 USD/t; the open marker on the central line is the point at which it leaves the voluntary anchor band, at 100 USD/MWh. Willingness-to-pay anchors are drawn in separate colour families: the amber band is the modelled voluntary durable-removal anchor (300 to 600 USD/t) and the amber dotted line the upper end of observed transactions (about 800 USD/t); the red dashed line is the 45Q credit (180 USD/t) and the red dotted line twice that value; the purple band is the EU ETS allowance (65 to 90 USD/t). (b) Distribution of LCOD across the 200,000-draw joint uncertainty analysis of Section 3.5, on the same cost axis, annotated with the share of draws falling below each anchor. The figure supports one principal conclusion: the modelled cost distribution overlaps the voluntary anchors but lies wholly above both compliance anchors, so the binding constraint differs by market segment, being volume in the voluntary market and price in the compliance market. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

0.9, and a reader who regards ReFuelEU as satisfying that anchor for DAC would score C1 at 0.7 and obtain a DAC index of 0.66 with a Tier-1 sub-index of 0.79, against the 0.71 and 0.86 reported. The headline coding is retained at 0.9 on the ground that the mandated volumes before 2035 are small relative to the announced DAC pipeline, and that the obligation falls on fuel suppliers rather than on removals. That fossil point-source CO2 remains compliant in the interim, so no purchase obligation for atmospheric CO2 binds at any price inside the modelled cost range in the period assessed. That judgment is contestable; the alternative is reported here and in Supplementary Table S9, and neither value changes the cross-technology ordering. Second, and independently of the coding, the integrated pathway is a genuine mitigation route that the compound-failure account must accommodate: it addresses the demand node without addressing cost, and the framework's own logic therefore predicts that it relieves the structural gap only in combination with cost reduction in electrolysis and capture. The electrofuel premium over conventional jet fuel is currently larger than the ammonia premium modelled in Section 4.3, consistent with that prediction and providing a natural next test of it.

The market is thin, but it exists, and the qualification runs counter to this paper's case and should be stated in full. An operating voluntary market for durable removal trades at roughly 300–800 USD/t under multi-year corporate offtake agreements. This is wider at the top than the 300 to 600 USD/t anchor band used in the cost comparison above, which is set to the range in which contracted volume is concentrated rather than to the full spread of observed prices, with DAC increasingly preferred within it for the permanence of geological storage; across the joint draws of Section 3.5, modelled LCOD falls below 800 USD/t in 84% of cases, so at the upper end of that range the technology is not priced out. A revenue channel of this kind is a demand condition that clean hydrogen never had at any point in its announcement cycle, and an account that treats DAC as simply lacking demand would be wrong. What the market lacks is depth rather than price: demand for high-durability credits depends on voluntary offsetting, and through 2025 it rested heavily on a single dominant buyer [17]. Climeworks, the best-funded DAC company, cut roughly 22% of its workforce in May 2025 while its plants ran far below nameplate capacity, capturing on the order of 1100 t against roughly 380,000 t contracted, with its 36,000-t Mammoth plant removing about 805 t in ten months [17,18]. Policy support is also fragile: in October 2025 the U.S. Department of Energy terminated more than USD 7.5 billion across 223 awards, cutting roughly ten DAC hub awards, before confirming in April 2026 that the two flagship megaton hubs would retain about USD 1.2 billion combined, even as the FY2027 budget request proposed rescinding USD 2.3 billion of the USD 3.5 billion programme [12,19]. This whipsaw is the policy-uncertainty cause (C2) that paralysed hydrogen: capital cannot price a subsidy that may vanish and reappear within a budget cycle. That the United States appears here as the leading instance of policy uncertainty and in Section 5.1 as the jurisdiction with the highest hydrogen delivery rate is not a contradiction, and the two characterisations are worth reconciling explicitly. They concern different instruments and periods: the delivery-rate comparison reflects the demand-pull and offtake provisions in force through 2024, whereas the whipsaw described here is a 2025-to-2026 reversal of discretionary appropriated support. The pairing that produced the delivery rate is precisely what the DAC programme has since lost, which is the point of the comparison, not an inconsistency in it. With project financing costs plausibly above 12%, on the benchmark basis and with the caveats set out in Section 3.2(i), and only 27 plants operating worldwide against roughly 130 announced [11], confirmed capacity is likely to track a small fraction of the pipeline through 2030.

Fig. 6 assembles these comparisons, and its two panels answer different questions. Panel (a) plots modelled cost against the industrial electricity price and separates two colour families: blues are modelled cost, with the central and first-of-a-kind cases as solid lines, the shaded band spanning the range of the non-energy cost A at the central energy intensity, and the dashed lines the full envelope once ηE is varied as well; amber, red and purple are willingness-to-pay anchors, drawn as bands where the anchor is a range and as lines where it is a single value. Panel (b) places the joint uncertainty distribution of Section 3.5 on the same cost axis, so that the position of the modelled distribution can be read directly against each anchor. The principal conclusion the figure supports is the asymmetry between the two anchor families. The cost distribution overlaps the voluntary anchors substantially, with 42% of draws below 600 USD/t and 84% below 800, so the voluntary market is a price the technology can sometimes meet; it lies entirely above both compliance anchors, with no draw below the 45Q credit and 0.9% within a factor of two of it, so no compliance mechanism now in force is a price it can meet at all. A secondary reading concerns which parameter drives the result: the vertical extent of the cost band at any given electricity price, 650 USD/t, exceeds the 160 USD/t by which the central case moves across the entire 40 to 120 USD/MWh range, which is the graphical form of the statement in Section 3.2(i) that A rather than the energy term dominates the gap.

4.3. *Green ammonia: the compliance gap*

Green ammonia inherits hydrogen's failure structure because it is synthesised from green hydrogen. Its demand gap is a compliance gap (Fig. 7). The breakeven penalty, the carbon price at which switching from conventional fuel becomes economic, is 367 USD/t CO_2 in the most favourable case ($E = 40$ USD/MWh, $CF = 0.60$), 528 centrally and 682 at European conditions, against a green premium of 1170 to 2180 USD/t MGOe. The International Maritime Organization's remedial-unit prices for 2028 to 2030 are 100 and 380 USD/t CO_2-equivalent. The lower tier lies far below breakeven under every parameter combination tested; reaching it would require green hydrogen at 1.36 USD/kg, which is not attainable at any electricity price above about 2 USD/MWh in Eq. (6). The upper tier of 380 marginally *exceeds* the most favourable breakeven of 367, so fuel-switching is economic in that corner, and the earlier claim that the penalties sit below breakeven across the whole range is corrected accordingly. The corner is narrow: it requires firm renewable

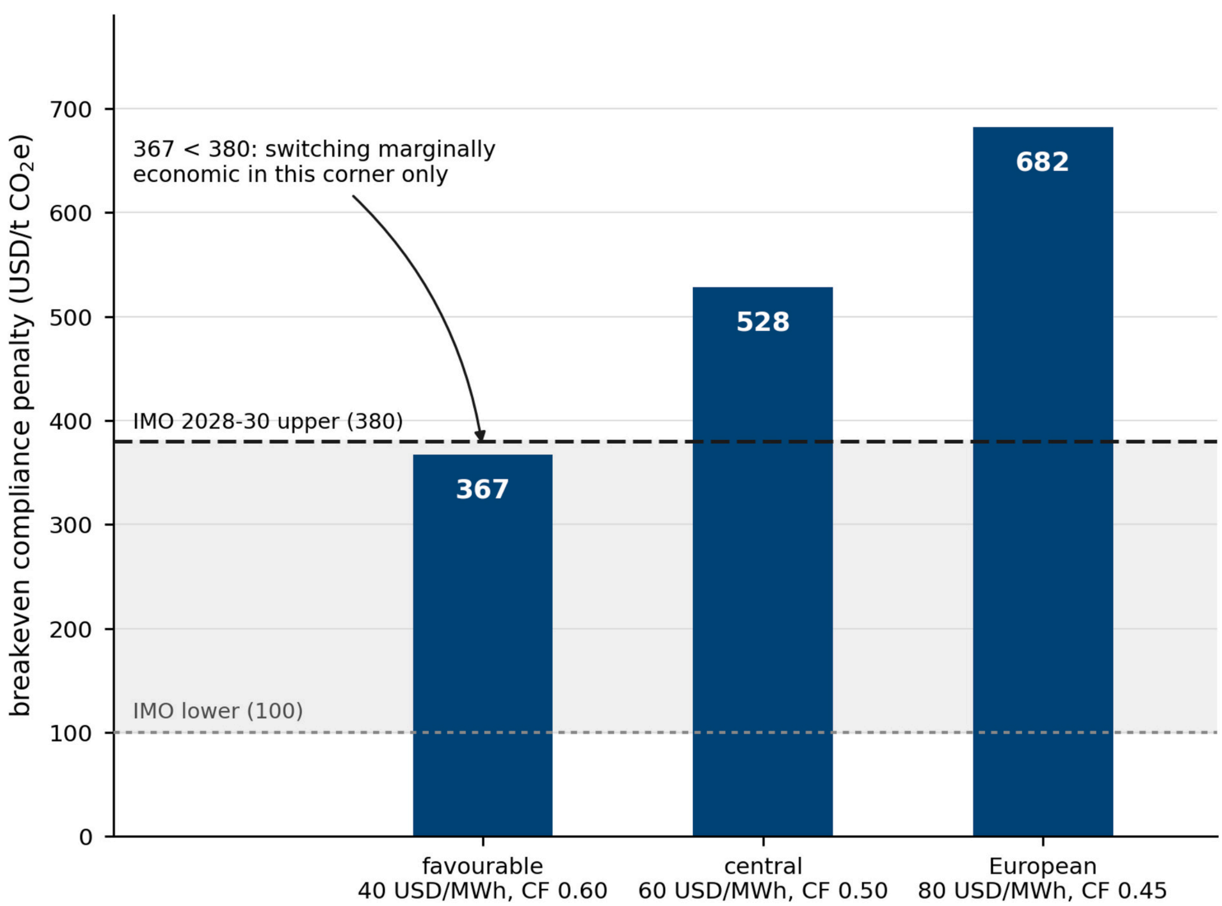

Fig. 7. Green-ammonia breakeven compliance penalty by scenario against the IMO remedial-unit prices; shading marks penalties below breakeven. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

electricity at or below about 42 USD/MWh at a 60% capacity factor, and in the joint uncertainty analysis of Section 3.5 the breakeven penalty falls below 380 USD/t CO2 in 4.3% of draws (median 564; 90% interval 385 to 752). The operative conclusion is therefore narrower but better supported than the original. Under the announced remedial-unit prices, fuel-switching is uneconomic for the overwhelming majority of the world fleet and becomes marginally economic only at the very best-resourced production sites, so paying the penalty remains the rational default well into the 2030s [14,21].

The demand signal is also not yet binding. The IMO Net-Zero Framework, approved at MEPC 83 in April 2025, would apply mandatory intensity targets to ships above 5000 gross tonnage. However, adoption was adjourned for twelve months in October 2025 (57 to 49, 21 abstentions), and following MEPC 84 in April to May 2026, the decision is now scheduled for a resumed extraordinary session on 4 December 2026, with entry into force expected around spring 2028 [21–24]. The project record reflects the gap: Australia's export pipeline has seen cancellations and indefinite holds (CQ-H2, Origin Energy's Hunter Valley project, Provaris, Desert Bloom), EverWind's Nova Scotia offtake to Germany evaporated, and Fortescue cancelled its Arizona and Gladstone projects citing high energy costs and absent demand [25,26]. Where projects advance, they restructure offtake around the gap rather than closing it: at NEOM, Air Products remains sole offtaker of up to 1.2 Mtpa, with Yara commercialising the remainder to de-risk the project [27]. This is worth reading as more than a workaround. Vertical integration across the electrolysis, synthesis and offtake steps, and diversification across the fertiliser and marine markets, insulate a producer from the merchant green premium that the standalone breakeven of Fig. 7 assumes, and the breakeven should therefore be understood as applying to a merchant transaction between an unaffiliated producer and shipowner. It is the right object for assessing whether a compliance mechanism will move the marginal cargo, which is the question the Net-Zero Framework poses. However, it overstates the exposure of integrated producers and understates the extent to which the industry may proceed through integration rather than through market formation. That the surviving projects are disproportionately integrated is consistent with this reading, and it qualifies the claim that green ammonia inherits hydrogen's failure structure: it inherits the cost structure, but the industrial configurations available to it differ.

4.4. Long-duration storage: the incumbent-capture squeeze

LDES is the positive control, and the levelised-cost model explains why demand alone did not build a frontier-LDES industry (Fig. 8). Under the central case of a common 40 USD/MWh charging price, lithium-ion is the cheapest option below about 10 h, and pumped hydro is the cheapest above it. None of the three frontier chemistries is least-cost at any duration in the range the Great Britain scheme procured (mostly 8 to 24 h): each is undercut by lithium-ion at short duration and by pumped hydro at long. This result holds across the full 20 to 60 USD/MWh charging-price sensitivity, at every point of which the least-cost technology across 8 to 24 h is either lithium-ion or pumped hydro (Supplementary Fig. S2), and it also holds under an equal-cycle counterfactual.

Under technology-differentiated charging, however, the result does not hold unqualified, and the qualification is worth stating precisely because it bears on the policy conclusion. If low-efficiency, infrequently cycling assets can systematically charge at 15 to 30 USD/MWh on surplus energy. In contrast, lithium-ion charges at 40, vanadium redox flow becomes the least-cost option at 8 h, by a margin of about 4 USD/MWh, and iron-air closes to within about 10% of pumped hydro at 24 h. The frontier chemistries therefore have a competitive window. However, it is narrow, sits at the short end of the procured range where lithium-ion is strongest, and exists only under a charging-price advantage that the cap-and-floor mechanism does not confer and that no bidder can underwrite ex ante. This is a sharper statement of the squeeze than the original

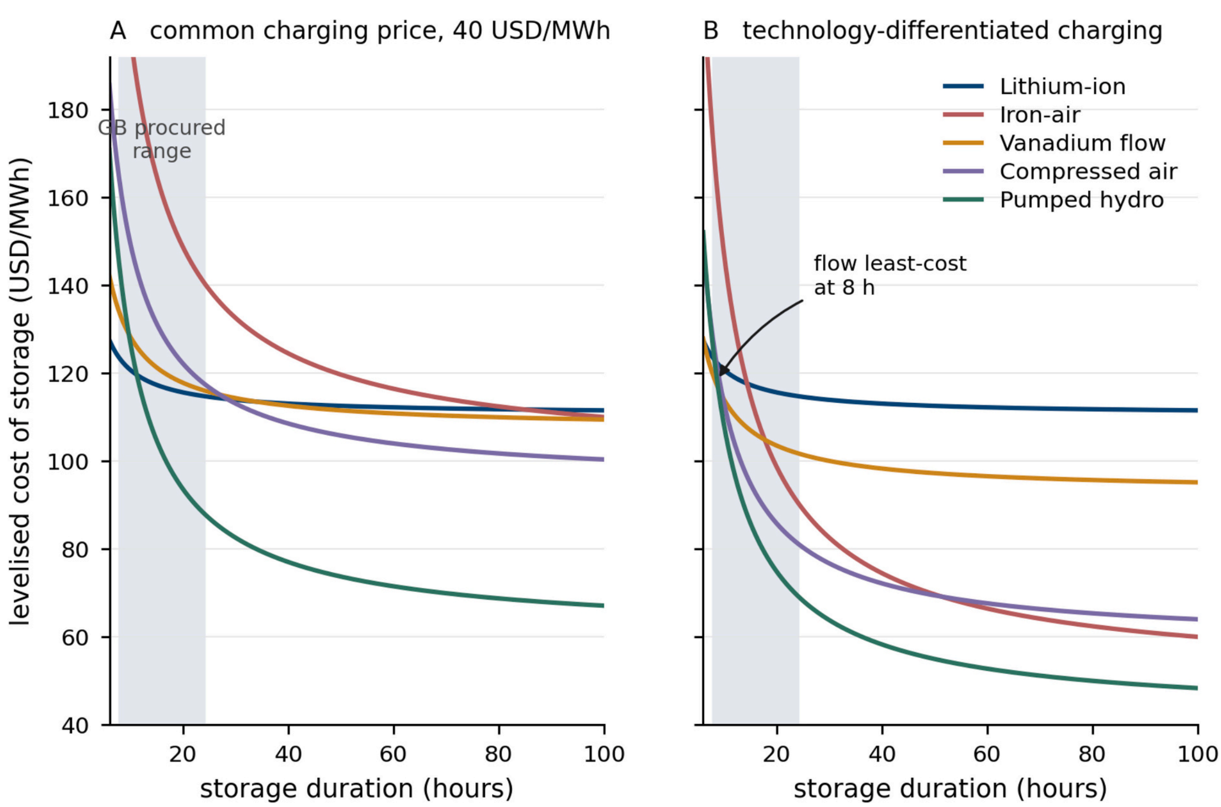

Fig. 8. Levelised cost of storage versus duration for lithium-ion, iron-air, vanadium redox flow, compressed air and pumped hydro; shading marks the range procured by the Great Britain scheme.

single-composite treatment allowed: the frontier is not uncompetitive everywhere, but it is uncompetitive under the revenue and dispatch conditions the scheme actually created.

The squeeze is structural rather than transitional. Even as global long-duration deployments rose about 49% in 2025, falling lithium-ion prices are making the incumbent competitive at ever-longer durations, compressing the window described above and squeezing the emerging chemistries the scheme was meant to support [28].

Two limits of the control should be acknowledged, and both cut in the same direction. First, the technology that responded to the mechanism is largely mature: pumped storage hydro has been built for a century, and the 171 bids came from a sector ready to build. The case therefore demonstrates that a revenue mechanism mobilises capital where a deployable technology already exists, which is a weaker claim than that a comparable mechanism would mobilise capital for DAC, where no such technology is yet deployable at scale. In this specific respect, the positive control is the case least like clean hydrogen in its announcement phase, and inferences from it to DAC should be drawn with that in mind. Second, a single application window in a single jurisdiction admits alternative explanations for the incumbent-heavy outcome that this analysis does not exclude: the eligibility and assessment criteria may themselves favour proven technologies, the shortlist is 'minded-to' rather than final, and Great Britain's unusually large consented pumped-hydro pipeline is a national endowment rather than a general condition. What the case establishes robustly is the narrower and still consequential point that resolving demand alone did not direct capital to the frontier; what it does not establish is the counterfactual scheme design that would have.

Great Britain provides the natural experiment. After nearly four decades of near-zero deployment under energy-only arrangements (a standing fleet of about 2.8 GW of pumped hydro), the government confirmed a cap-and-floor scheme in October 2024; its first window drew 171 bids (52.6 GW), of which 77 projects (28.7 GW) cleared eligibility screening in September 2025, 73 proceeding to full assessment. On 26 June 2026, Ofgem published a 'minded-to' shortlist of 16 projects totalling 7.6 GW and 136.9 GWh [29–31]. The demand mechanism worked, but the awarded capacity flowed overwhelmingly to the incumbents. The shortlist spans four technologies, yet pumped hydro alone accounts for roughly two-thirds (64%) of the shortlisted energy and lithium-ion, admitted only after industry pressure and over objections that the scheme distorts the market, takes most of the remainder; only two smaller projects use non-incumbent chemistries, one compressed-air and one combining vanadium redox flow with zinc-hybrid cells [31].

Meanwhile, the frontier firms failed on cost and finance: ESS Inc. posted a 2025 net loss of about USD 63 million on USD 1.6 million in revenue and is out of cash; EOS is missing revenue; and Morrow filed for bankruptcy, the third Nordic maker to fall after Northvolt and Freyr [32,33,35]. The one winner, Form Energy, grew its pipeline by about 375% on a roughly USD 1 billion Google agreement underpinned by tax credits and a single anchor buyer, mirroring DAC's reliance on a single purchaser [34]. Revenue certainty determined that storage would be built; cost and finance determined which storage, and the answer was overwhelmingly the incumbent.

5. Design implications and limitations

The results point to a small number of design choices rather than more capital. Three propositions follow (Sections 5.1–5.3). Section 5.4 separates the three distinct senses in which a technology may be said to be ready; Section 5.5 states the analytical boundaries within which the three cases are comparable. Section 5.6 sets out the limitations of the analysis and the work required to overcome them.

5.1. *Demand-side mechanisms must precede supply-side capital, but pairing determines whether an industry becomes self-sustaining*

The counterfactual quantifies the size of the effect the framework implies: resolving demand alone lowers the DAC and green-ammonia indices by under 0.10, whereas addressing demand and cost together lowers them by about 0.20 (Table 4). As Sections 3.1.4 and 4.1 make explicit, that contrast is a property of the index rather than evidence for the proposition, and it is the field evidence that carries the argument. That evidence points the same way. Production subsidies cannot manufacture a buyer, and the two policy sequences diverge sharply (Fig. 3): the supply-led path stalls at FID when assumed demand fails, while the demand-first path elicits a developer response, as Great Britain's scheme did. The clean-hydrogen record is consistent with the sequencing point (Japan about 17% and South Korea about 11% delivery, the United States about 25%, against 3 to 6% for supply-led economies). However, it is weaker evidence than a bare comparison of rates suggests, for two reasons that should be stated. First, these are delivery rates and therefore require a denominator of announced capacity, which the cancellation record [8] does not contain; they are taken from the agency and industry reviews [36,37], and the cancellation record is cited here only for the cause attribution, not for the rates. Second, and more substantively, the jurisdictions differ in more than policy sequence, and the United States is not a demand-first case at all but a hybrid one, combining binding industrial offtake commitments with large-scale supply-side incentives, notably the Section 45 V production credit and federal loan guarantees. Its 25% cannot therefore be attributed solely to demand pull, and reading it that way would overstate the argument. The comparison is offered as consistent with the proposition rather than as identification of it. The LDES case supplies the qualifier that the index cannot, because their demand was resolved by policy while cost and finance were not, and capital flowed to whichever technology was already cheapest under the dispatch conditions the scheme created (Section 4.4). Unless a demand mechanism is paired with cost reduction and de-risking finance, it subsidises the incumbent.

5.2. *Commitments should be screened for compound risk before they are made*

The CDRI operationalises exactly such a screen: it would have flagged DAC and green ammonia as high-risk on structural grounds in 2025 to 2026 (Tier-1 sub-index 0.86) and LDES as materially lower (0.47), while still flagging the frontier-LDES chemistries through the cost and finance causes. Single-cause assessment cannot do this. Neither can a maturity ladder such as the TRL or CRI [38,39], which position a technology on an ordinal scale without an explicit aggregation rule and therefore cannot distinguish a technology that clears every tier from one that clears only the demand tier, nor compute the effect of resolving one barrier against two. Compensatory MCDM methods such as AHP and TOPSIS [45,46] can rank technologies based on the same eight causes, and Section 3.1.3 shows they reproduce the ordering reported here; however, they cannot restrict substitution between causes or support node-resolution counterfactuals. The claim made for the CDRI is bounded accordingly: it is a screening instrument for compound risk with a stated aggregation axiom, not a general-purpose readiness measure, and it is complementary to, not a replacement for, technical maturity assessment (Section 5.4).

Three practical constraints on using it as a screen should be stated, since an index that is silent on them invites arbitrary application. First, weighting. The index weights all eight causes equally, which is a deliberate default rather than a finding: no evidence exists for differential weights, and the cause-share data from the cancellation record cannot supply them, because shares measure how often a cause appears first and not how much it matters conditional on appearing. A weighted variant is trivial to compute within Eq. (5) and is provided in the Supplementary dataset, but any particular weighting would be an assertion.

Second, thresholds. No threshold value of the index is proposed, and none should be inferred, because the index is ordinal (Section 3.1.2); a screen built on it compares a candidate against scored reference technologies, of which clean hydrogen at its 2021 announcement peak is the natural benchmark, rather than against an absolute cut-off. Third, the response should be graduated rather than binary. The useful output of the screen is not a proceed or defer decision but the identification of *which* causes are binding and therefore which instruments would have to be paired: a high index driven by demand and cost calls for a revenue mechanism paired with cost support, whereas the same index driven by permitting and grid access calls for consenting and connection reform, and the two are not interchangeable. Used that way, the index informs instrument choice; used as a gate, it would carry more weight than its construction supports.

5.3. Policy timelines must match cause latency

The three latencies used here are order-of-magnitude estimates drawn from the empirical record rather than model outputs, and their basis should be stated. The 1 to 3 year figure for rulemaking and permitting reform is taken from the observed interval between instrument proposal and entry into force in the cases assembled in this paper and in [8], of which the IMO Net-Zero Framework is representative: approval at MEPC 83 in April 2025, an adoption decision scheduled for December 2026 and entry into force expected in spring 2028. The 5- to 10-year cost-reduction figure follows from the learning analysis below. The 10 to 15 year figure for demand-market development is the least well grounded of the three, resting on the diffusion intervals reported in the transitions literature [2,3] and on the hydrogen record, and it is offered as an indicative planning horizon rather than an estimate; Section 5.6 identifies its empirical determination as outstanding work.

The learning requirement can be stated more precisely than a latency band. Holding electricity at 60 USD/MWh, so that central LCOD is 520 USD/t, three targets should be distinguished, because the learning requirement is strongly non-linear in the target and an earlier version of this paragraph reported only the middle one, having misdescribed 300 USD/t as the mid-point of the 300 to 600 USD/t voluntary range rather than its lower bound. Reaching the actual mid-point of that band, 450 USD/t, requires the annualised non-energy cost A to fall from 400 to 330 USD/t, a 17.5% reduction, which is 1.2 capacity doublings at a 15% learning rate or 0.9 at 20%, on the order of a two- to two-and-a-half-fold expansion of installed capacity. Reaching the lower bound of the band, 300 USD/t, requires A to fall to 180 USD/t, a 55% reduction, which is 4.9 doublings at 15% or 3.6 at 20%, on the order of a twelve- to thirty-fold expansion. Reaching the 45Q credit at 180 USD/t requires an 85% reduction in A, some 8.5 to 11.7 doublings, which no plausible deployment trajectory delivers before the late 2030s. The three figures should be read together, and the correction matters for the interpretation: modest learning reaches the middle of the voluntary market, deep learning is needed to reach its lower bound, and the compliance anchors are out of reach on any near-term trajectory. That is the cost-side counterpart of the finding in Section 4.2 that the constraint at the top of the voluntary market is volume rather than price. For green ammonia, the binding requirement is not learning in the synthesis loop but the delivered price of firm renewable electricity: from Eq. (6), the upper IMO remedial-unit price of 380 USD/t CO_2 is reached only at or below about 42 USD/MWh at a 60% capacity factor, and the 100 USD/t tier is out of reach at any electricity price above about 2 USD/MWh (Section 4.3). These thresholds are the quantitative form of the pairing argument: they identify what cost-side action would have to accompany the demand-side instruments now being adopted for those instruments to bind. They also make clear that a single latency band should not be applied across technologies with different cost structures. DAC is modular, built from replicated contactor and regeneration units, and modular technologies have historically shown faster learning than capital-intensive assets whose cost reduction depends on site-specific engineering and scale [10,55]; that is why the DAC calculation above is run at 15% and 20% learning rates. Haber-Bosch synthesis, pumped storage hydro and compressed air are not modular; civil works and site conditions dominate their unit costs, and learning rates in the 5 to 10% range are the appropriate assumption for them, which is one reason the ammonia threshold above is expressed in terms of the delivered electricity price rather than of synthesis learning. The starting points differ as sharply as the rates: DAC begins from a far higher baseline cost than any of the storage technologies. Evaluating these technologies against a common fixed timeline would therefore be misleading, and the latencies in the preceding paragraph should be read as indicative planning horizons that require technology-specific adjustment, not as benchmarks.

For DAC and green ammonia, demand-side mechanisms activated in 2025 to 2026 will therefore not generate traction before the mid-2030s, so deferring demand-side action while expanding supply-side support replicates the hydrogen timeline mismatch. LDES shows both an alternative route to triggering investment and, equally, that the cost and finance latencies must be worked in parallel, or the demand mechanism will reward whichever technology has already descended its cost curve.

5.4. Engineering maturity, economic competitiveness and commercial deployment are distinct

The framework deliberately foregrounds market, policy, and finance barriers, which risks conflating three distinct properties that a technology may or may not possess. They should be separated explicitly. *Engineering maturity* is the demonstrated ability to build and operate the technology at the required scale and reliability; it is what the TRL scale measures. *Economic competitiveness* is the relation between the technology's delivered cost and the value it can capture; it is what the cost models of Section 3.2 measure. *Commercial deployment* is the realised commitment of capital at FID, which requires both of the above and, in addition, a buyer, a bankable revenue mechanism, consent and grid access; it is what the CDRI is designed to predict. The three do not move together, and the paper's central claim is precisely that the first does not imply the third.

The three technologies occupy different positions on the first axis, and this deserves more weight than the analysis has so far given it. Pumped hydro and lithium-ion are mature by any measure. Green ammonia synthesis is industrially mature on the Haber-Bosch side and immature on the electrolysis side at gigawatt scale, and immature again in marine propulsion, where toxicity handling, ammonia slip, N2O formation, pilot-fuel requirements and crew competency remain open. DAC is the least mature: the operating record in Section 4.2, of roughly 805 t removed in ten months at a 36,000-t nameplate plant, is an engineering and operating-performance shortfall, not a market failure, and it is coded in the index as the execution cause C8. The frontier storage chemistries sit between: iron-air and zinc-hybrid systems have limited multi-year field operating records, and vanadium flow systems face electrolyte cost and supply issues rather than performance questions.

Two consequences follow. First, the CDRI should be read alongside, not instead of, a technical maturity assessment; the two answer different questions, and a technology can be TRL 9 and still score highly on the CDRI, which is the situation this paper argues obtains for green ammonia production. Second, part of what the index attributes to structural market failure at DAC is properly attributed to operating performance, because a plant delivering only a fraction of its nameplate capacity cannot even serve the demand that exists. The C8 coding captures this, but its Tier-3 position, inherited from the hydrogen cancellation record where execution accounted for only 4% of cancellations, may understate its weight for a technology as immature as DAC. Whether the tier structure derived from hydrogen transfers to less mature technologies is an open question and is listed in Section 5.6.

This three-way distinction is not novel, and the framework should be positioned against the existing instrument for it. The United States Department of Energy's Adoption Readiness Level framework was developed precisely to complement the TRL scale by assessing the non-technical conditions for deployment, and organises them under four headings: value proposition, market acceptance, resource maturity, and license to operate [53]. The eight causes used here map onto that structure without residue: demand and production cost fall under value proposition; policy uncertainty and permitting under license to operate; finance, electricity access, and supply chain under resource maturity; and execution performance under market acceptance in its operational-track-record sense. The mapping is worth making explicit for two reasons. It provides independent corroboration that the eight causes span the relevant space rather than being an ad hoc list, which is a check the single-analyst derivation of Section 3.1.1 cannot supply on its own. Moreover, it locates the contribution precisely: the ARL framework assesses each dimension separately and does not aggregate, so it cannot express compounding across dimensions or support the counterfactuals of Section 3.1.4, which is what the CDRI adds; conversely, the ARL framework carries institutional vetting and cross-sector application that this index does not. A practitioner would reasonably use the ARL structure to organise evidence and the CDRI to aggregate it.

5.5. Analytical boundaries, comparability and transferability

The three applications are not drawn at identical analytical boundaries, and the claims made from them should be read accordingly. DAC is assessed as a global technology category against international policy and voluntary-market anchors. Green ammonia is assessed in a single end-use, marine fuel under the IMO framework, chosen because it is the only end-use with an approaching compliance obligation and therefore the only one in which the demand cause can be scored against a documented instrument; the conclusions do not transfer without modification to fertiliser feedstock, where a large merchant market already exists and the relevant gap is a green premium rather than an absent buyer, nor to hydrogen carriage or power generation, where the competing reference technology differs. LDES is assessed against one jurisdiction's procurement scheme, because that scheme is the only instance in which the demand cause has been resolved by design and an outcome observed.

A prior question is whether the three technologies are comparable to clean hydrogen at all, and the answer is that they are comparable in market position but not in technical maturity, which bounds the inference. On the technology-readiness dimension, using the IEA's clean-energy technology classification [54], the three do not align: pumped storage hydro and lithium-ion are commercially mature; ammonia divides, with synthesis mature and marine propulsion at demonstration; and DAC is the least mature and comprises several distinct routes at different stages, which this paper treats as one and which Section 3.2(i) identifies as a simplification. Clean hydrogen in its announcement phase was closer to the ammonia case than to either extreme. What the four do share, and what the framework actually requires of them, is a market position: a large announced pipeline, capital-intensive assets with long payback, an absent or thin buyer at achievable prices, and a dependence on policy for revenue. The comparison is therefore drawn on adoption readiness rather than technology readiness, in the sense of Section 5.4, and the paper's claim is correspondingly restricted: the compound-failure architecture is proposed as a transferable account of *why capital does not commit*, not as a claim that the four technologies are at similar stages of development. It also follows that the inflexion-point framing used in Section 1 should be read as a statement about announcement and commitment dynamics rather than about technical maturity, and the text has been adjusted accordingly.

The LDES case introduces a further ambiguity about the unit of analysis that should be made explicit. The CDRI of 0.44 is scored for the LDES category as a whole, including commercially established pumped hydro and lithium-ion, whereas the commercial-failure evidence in Section 4.4 concerns frontier chemistries and the firms developing them. The two are not the same object, and the reconciliation is the point of the case rather than a defect in it: the category-level index is low precisely because demand and, for the incumbents, cost are resolved, while the frontier-level failure runs through the cost and finance causes that the category-level score does not isolate. A frontier-only sub-index, scored separately for the emerging chemistries, is reported in the Supplementary dataset and is materially higher than the category score; the main-text value should be read as the category score throughout.

The three cost models likewise quantify different objects: a removal cost, a breakeven compliance penalty and a levelised storage cost, and they are not numerically commensurable with one another or with the index. The claim that the technologies are placed on a common quantitative footing should therefore be understood in a specific and limited sense. Each model is constructed to express the same underlying quantity: the distance between what the technology costs and what the best available buyer or compliance mechanism will pay, in the units set by that particular market. The index and the models are structurally linked, in that the models test whether the cost cause C3 is correctly coded as binding, rather than through a numerical mapping from cost gap to index score. Constructing such a mapping would require calibration data that do not exist.

With three heterogeneous cases and a single coder, no claim of general transferability across energy technologies is warranted. The claim made in Section 1 is correspondingly narrowed: the CDRI is offered as a transferable *instrument*, in the sense that its rubric, aggregation rule and uncertainty treatment are fully specified and reproducible for application to other technologies, not as an instrument whose *calibration* has been demonstrated to transfer. The latter requires the cohort study described below.

5.6. Limitations

Five limitations bound the conclusions. First, and most seriously, all cause-level scores were assigned by a single analyst. The behaviourally anchored rubric of Table 1 and the score-level provenance in Supplementary Table S1 make the coding auditable and contestable. The Monte Carlo perturbs every score by up to one full rubric level. However, neither is a substitute for independent coding: the perturbation propagates an assumed numerical uncertainty around scores that may themselves be systematically biased. No inter-coder reliability statistic can be reported. The structured expert judgment protocol specified in Supplementary Table S7, in which multiple calibrated experts score the rubric and are performance-weighted on calibration questions drawn from the resolved hydrogen record [40,41], is the appropriate remedy and is set out for future application rather than executed here.

Second, the framework's empirical foundation is the clean-hydrogen cancellation record assembled by the present author [8]. That dataset is openly archived and independently auditable. However, it is a self-authored data resource rather than an independently peer-reviewed methodological study, and the framework's three-tier structure, its causal-primacy hierarchy and the cause-share figures quoted in Section 2 all derive from it. The retrodictive exercise of Section 3.4 is a consistency check against that same record and cannot serve as external validation.

Third, the index is ordinal in construction, and its levels have no absolute meaning (Section 3.1.2). Fourth, the cost models are reduced-form, and their boundaries are stated in Section 3.2; the DAC model in particular collapses thermal and electrical energy into a single electricity-equivalent intensity and does not resolve process configuration or heat integration, and a configuration-resolved successor distinguishing solid-sorbent and liquid-solvent pathways with an explicit heat price is the priority extension. Fifth, the tier structure and cause weights are inherited from a single technology's cancellation record, and their transferability to technologies at different maturity is untested (Sections 5.4 and 5.5). Sixth, and most consequentially for the framework's central claim, the underlying record codes one primary cause per event and therefore cannot evidence causal interaction (Section 2);

recoding the archived cancellations for all operative causes, and reporting the joint distribution, is the work that would convert the conjunctive architecture from an imported axiom into a finding. It is a prerequisite for the weighted variant of the index discussed in Section 5.2.

The decisive next step for the instrument is neither a further robustness treatment nor an additional case. It is a genuinely ex-ante test: scoring a cohort of capital-intensive climate technologies today, pre-registering the scores and the outcome measure, and evaluating realised FID and cancellation outcomes at a fixed horizon. Only that design can establish whether the CDRI predicts, as distinct from describing coherently and retrodicting consistently. The rubric, aggregation rule, uncertainty treatment, and elicitation protocol are specified here in the form a study of this type would require.

6. Conclusions

A reproducible compound deployment risk index and three cost-gap models place the three technologies on a common footing, in the specific sense set out in Section 5.5. Each model expresses the distance between what a technology costs and what its best available buyer or compliance mechanism will pay, in the units in which that market decides. DAC and green ammonia score high (0.71 and 0.73) because costs compound the absence of structural demand. The modelled DAC removal cost exceeds the compliance anchors for every parameter combination tested. It clears the top of the voluntary durable-removal range only where industrial electricity is below about 100 USD/MWh, so that at the high-price end of the voluntary market the binding constraint is volume rather than price. The green-ammonia breakeven penalty (370 to 680 USD/t CO2) exceeds the IMO remedial-unit prices (100 to 380) except in a narrow best-site corner reachable only with firm renewable power at or below about 42 USD/MWh. LDES scores markedly lower (0.44) because its demand cause has been addressed by design, and no frontier chemistry is least-cost at any procured duration under a common charging price, though a flow chemistry becomes marginally least-cost at eight hours if low-efficiency assets can systematically charge on surplus energy; the Great Britain awards accordingly flowed overwhelmingly to pumped hydro and lithium-ion.

The counterfactual quantifies what the framework's aggregation axiom implies at the observed scores, and its ordering is invariant across the whole power-mean family of aggregators, across the rubric mappings tested, under the minimum-based variant of the aggregation rule, and under coding uncertainty of one full rubric level with correlated coder error. It is not, however, evidence for the proposition it quantifies, since a larger response to two changed inputs than to one follows arithmetically from the rule. The evidence lies in the cost-gap models, in the retrodictive consistency check against the clean-hydrogen record, and above all in the LDES case, where policy resolved demand and the market still selected the incumbent. These point the same way: demand-side design is necessary but not sufficient. Sequencing a revenue mechanism ahead of capital determines whether investment happens; pairing it with cost reduction and de-risking finance determines whether a self-sustaining industry follows. What this paper establishes is a diagnostic that is explicit about its own axioms and boundaries, and that is reproducible on other technologies; what it does not yet establish is predictive performance, which requires the pre-registered cohort study set out in Section 5.6. Whether demand-side policy is paired rather than deployed alone remains a policy choice, not a technical one.

Declaration of generative AI and AI-assisted technologies in the writing process

During the preparation of this work, the author used Claude for proofreading and language polishing. After using this tool, the author carefully reviewed, edited, and verified all content generated to ensure accuracy and appropriateness. The author take full responsibility for the content of the published article.

CRediT authorship contribution statement

Sinan Küfeoğlu: Writing – review & editing, Writing – original draft, Methodology, Investigation, Formal analysis, Data curation, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.apenergy.2026.128871.

Data availability

All rubric scores with anchor-condition provenance and full five-level anchor sets (Table S1), cost-model parameters (Table S2), index computations, aggregation-rule, rubric-mapping and dependence robustness tables (S3, S3b, S4, S5), correlated Monte-Carlo results (seed 42), the retrodiction coding (Table S6), the SEJ elicitation protocol (S7), the joint techno-economic Monte-Carlo results (Table S8), the frontier-only LDES sub-index, the equal-cycle and technology-differentiated charging cases, the analysis script, and cost-gap models including the charging-price sensitivity (Fig. S2) are provided in the Supplementary materials accompanying this article and are fully reproducible on Zenodo (https://doi.org/10.5281/zenodo.21919928). The clean-hydrogen cancellation dataset underlying the framework is openly archived on Zenodo (https://doi.org/10.5281/zenodo.20054872). All other data are drawn from the publicly available agency, regulatory, government and company sources cited in the References.

References

[1] Jaffe AB, Newell RG, Stavins RN. A tale of two market failures: technology and environmental policy. Ecol. Econ. 2005;54(2–3):164–74. https://doi.org/10.1016/j.ecolecon.2004.12.027.
[2] Nemet GF. Demand-pull, technology-push, and government-led incentives for non-incremental technical change. Res. Policy 2009;38(5):700–9. https://doi.org/10.1016/j.respol.2009.01.004.
[3] Geels FW. Technological transitions as evolutionary reconfiguration processes: a multi-level perspective and a case-study. Res. Policy 2002;31(8–9):1257–74. https://doi.org/10.1016/S0048-7333(02)00062-8.
[4] Bergek A, Jacobsson S, Carlsson B, Lindmark S, Rickne A. Analyzing the functional dynamics of technological innovation systems: a scheme of analysis. Res. Policy 2008;37(3):407–29. https://doi.org/10.1016/j.respol.2007.12.003.
[5] Hekkert MP, Suurs RAA, Negro SO, Kuhlmann S, Smits REHM. Functions of innovation systems: a new approach for analysing technological change. Technol Forecast Soc Change 2007;74(4):413–32. https://doi.org/10.1016/j.techfore.2006.03.002.
[6] Hartley PR, Medlock III KB. The valley of death for new energy technologies. Energy J. 2017;38(3):33–61. https://doi.org/10.5547/01956574.38.3.phar.
[7] Joskow PL. Capacity payments in imperfect electricity markets: need and design. Util. Policy 2008;16(3):159–70. https://doi.org/10.1016/j.jup.2007.10.003.
[8] Küfeoğlu S. Eight root causes of clean hydrogen investment failure: supplementary dataset and levelised-cost-of-hydrogen model [dataset]. Zenodo; 2026. https://doi.org/10.5281/zenodo.20054872.
[9] International Energy Agency. Direct air capture: a key technology for net zero. Paris: OECD Publishing; 2022. https://doi.org/10.1787/bbd20707-en.
[10] Sievert K, Schmidt TS, Steffen B. Considering technology characteristics to project future costs of direct air capture. Joule 2024;8(4):979–99. https://doi.org/10.1016/j.joule.2024.02.005.
[11] International Energy Agency. Direct air capture. Energy system: carbon capture, utilisation and storage. Paris: IEA; 2025. https://www.iea.org/energy-system/carbon-capture-utilisation-and-storage/direct-air-capture.
[12] U.S. Department of Energy. Energy department announces termination of 223 projects, saving over $7.5 billion [press release]. Washington, DC: U.S. DOE; 2025. 2 October, https://www.energy.gov/articles/energy-department-announces-termination-223-projects-saving-over-75-billion.

[13] Internal Revenue Code § 45Q, as amended by the Inflation Reduction Act of 2022, Pub. L. No. 117–169. 2022 (credit of USD 180 per tonne CO2 for direct air capture with secure geological storage).
[14] DNV. Ammonia as a marine fuel: prospects and challenges. Maritime impact. Høvik: DNV; 2025. https://www.dnv.com/expert-story/maritime-impact/ammonia-as-a-marine-fuel-prospects-and-challenges/.
[15] Viswanathan V, Mongird K, Franks R, Li X, Sprenkle V, Baxter R. 2022 grid energy storage technology cost and performance assessment. PNNL-33283. Richland (WA): Pacific Northwest National Laboratory for the U.S. Department of Energy Energy Storage Grand Challenge; 2022.
[16] Pacific Northwest National Laboratory. Energy storage cost and performance database v2024. Richland (WA): PNNL; 2024. https://www.pnnl.gov/ESGC-cost-performance.
[17] Climeworks AG. Statement on organisational restructuring and workforce reduction [press release]. Zurich: Climeworks; 2025.
[18] Climeworks AG. Plant performance disclosures: cumulative capture versus contracted volumes and nameplate capacity, Orca and mammoth facilities. Zurich: Climeworks; 2025.
[19] U.S. Department of Energy. Award retention list submitted to Congress, April 2026, retaining approximately USD 1.2 billion for Project Cypress and the South Texas DAC Hub; and FY2027 congressional budget request, 3 April 2026, proposing rescission of USD 2.3 billion of the USD 3.5 billion Regional Direct Air Capture Hubs programme. Washington, DC: U.S. DOE; 2026.
[20] Lee M, Saygin D. Financing cost impacts on cost competitiveness of green hydrogen in emerging and developing economies. OECD environment working papers no. 227. Paris: OECD Publishing; 2023. https://doi.org/10.1787/15b16fc3-en.
[21] International Maritime Organization. Report of the marine environment protection committee on its eighty-third session. MEPC 83/17. London: IMO; 2025. . (IMO Net-Zero Framework: greenhouse-gas fuel-intensity standard for ships above 5000 GT; remedial-unit prices of USD 100 and USD 380 per tonne CO2eq for 2028–2030.).
[22] International Maritime Organization. Report of the marine environment protection committee on its second extraordinary session (MEPC/ES.2), 14-17 October 2025: adjournment of net-zero framework adoption by 57 votes to 49 with 21 abstentions. London: IMO; 2025.
[23] Global Maritime Forum. A guide to the IMO'S net-zero framework. Copenhagen: Global Maritime Forum; 2026. https://globalmaritimeforum.org/news/a-guide-to-the-imos-net-zero-framework/.
[24] DNV. IMO MEPC 84: Revisiting the net-zero framework. Høvik: DNV; 2026. https://www.dnv.com/news/2026/imo-mepc-84-revisiting-the-net-zero-framework/. [MEPC 84, 27 April to 1 May 2026; resumed extraordinary session and adoption decision scheduled for 4 December 2026.].
[25] Origin Energy Ltd. Provaris Energy Ltd; and other developers. Company and exchange (ASX) announcements on the withdrawal, suspension or cancellation of Australian and North American green-ammonia and green-hydrogen export projects (CQ-H2; Hunter Valley; Tiwi Islands; Desert Bloom; EverWind Fuels; NextEra; G2 Net-Zero), 2024–2025. 2024.
[26] Fortescue Ltd. Quarterly production report, June 2025 quarter [ASX announcement]. Perth: Fortescue Ltd; 2025 (Discontinuation of the Arizona Hydrogen and PEM50 Gladstone projects; pre-tax write-down of USD 150 million.).
[27] Air Products and Chemicals, Inc.; NEOM Green Hydrogen Company. Offtake arrangements for renewable ammonia: Air Products sole offtaker of up to 1.2 Mtpa, with Yara commercialising residual volumes [company statements]. 2025.
[28] Wood Mackenzie. Long-duration energy storage market outlook: deployments and lithium-ion competitiveness at longer durations. Edinburgh: Wood Mackenzie; 2026.
[29] Department for Energy Security and nNet Zero. Long duration electricity storage: cap and floor scheme, government response. London: DESNZ; 2024.
[30] Ofgem. Long duration electricity storage cap and floor, window 1: eligibility assessment and project assessment framework decision. OFG1164. London: Ofgem; 2025 [171 applications, 52.6 GW; 77 projects, 28.7 GW, eligible at screening; 73 proceeding to assessment.].
[31] Ofgem.. Long duration electricity storage window 1: minded-to decisions [consultation]. London: Ofgem; 2026 [16 projects, 7645 MW and 136.9 GWh, across pumped storage hydro, lithium-ion, compressed air and vanadium redox flow with zinc-hybrid cells.].
[32] ESS Tech, Inc.. Annual report on form 10-K for the fiscal year ended 31 December 2025. Washington, DC: U.S. Securities and Exchange Commission; 2026.
[33] Eos Energy Enterprises, Inc.. Annual report on form 10-K for the fiscal year ended 31 December 2025. Washington, DC: U.S. Securities and Exchange Commission; 2026.
[34] Form Energy, Inc. Xcel Energy Inc. Announcement of a 300 MW / 30 GWh, 100-hour iron-air multi-day storage deployment supporting a Google data centre at Pine Island, Minnesota [company statements]. 2026.
[35] Morrow Batteries ASA. The board of directors of morrow batteries ASA has resolved to file for bankruptcy proceedings [press release]. Arendal: Morrow Batteries; 2026.
[36] International Energy Agency. Global hydrogen review 2025. Paris: IEA; 2025. https://www.iea.org/reports/global-hydrogen-review-2025.
[37] Hydrogen Council, McKinsey & Company. Global hydrogen compass 2025. Brussels: Hydrogen Council; 2025. https://compass.hydrogencouncil.com/.
[38] Australian Renewable Energy Agency. Commercial readiness index for renewable energy sectors. Canberra: ARENA; 2014. https://arena.gov.au/assets/2014/02/Commercial-Readiness-Index.pdf.
[39] IEA-RETD. RE-CRI: commercial readiness index assessment - using the method as a tool in renewable energy policy design. Utrecht: IEA Renewable Energy Technology Deployment Technology Collaboration Programme; 2017.
[40] Cooke RM. Experts in uncertainty: opinion and subjective probability in science. New York: Oxford University Press; 1991.
[41] Colson AR, Cooke RM. Expert elicitation: using the classical model to validate experts' judgments. Rev. Environ. Econ. Policy 2018;12(1):113–32. https://doi.org/10.1093/reep/rex022.
[42] OECD, European Commission Joint Research Centre. Handbook on constructing composite indicators: methodology and user guide. Paris: OECD Publishing; 2008. https://doi.org/10.1787/9789264043466-en.
[43] Nelsen RB. An introduction to copulas. 2nd ed. New York: Springer; 2006.
[44] Bullen PS. Handbook of means and their inequalities. Dordrecht: Kluwer Academic Publishers; 2003.
[45] Hwang CL, Yoon K. Multiple attribute decision making: Methods and applications - a state-of-the-art survey. Lecture notes in economics and mathematical systems186. Berlin: Springer-Verlag; 1981.
[46] Saaty TL. The analytic hierarchy process: planning, priority setting, resource allocation. New York: McGraw-Hill; 1980.
[47] Keith DW, Holmes G, St. Angelo D, Heidel K. A process for capturing CO2 from the atmosphere. Joule 2018;2(8):1573–94. https://doi.org/10.1016/j.joule.2018.05.006.
[48] Fasihi M, Efimova O, Breyer C. Techno-economic assessment of CO2 direct air capture plants. J. Clean. Prod. 2019;224:957–80. https://doi.org/10.1016/j.jclepro.2019.03.086.
[49] International Maritime Organization. 2023 guidelines on life cycle GHG intensity of marine fuels (LCA guidelines). Resolution MEPC.376(80). London: IMO; 2023.
[50] Schmidt O, Melchior S, Hawkes A, Staffell I. Projecting the future levelized cost of electricity storage technologies. Joule 2019;3(1):81–100. https://doi.org/10.1016/j.joule.2018.12.008.
[51] Ku AY, Greig C, Larson E. Three strategies to revive teetering clean hydrogen dreams. Energy Res. Soc. Sci. 2024;113:103576. https://doi.org/10.1016/j.erss.2024.103576.
[52] European Union. Regulation (EU) 2023/2405 of the European Parliament and of the council of 18 October 2023 on ensuring a level playing field for sustainable air transport (ReFuelEU aviation). OJ L 2023/2405, 31.10.. 2023.
[53] U.S. Department of Energy, Office of Technology Transitions. Adoption readiness levels (ARL): a complement to technology readiness levels. Washington, DC: U.S. DOE; 2023.
[54] International Energy Agency. ETP clean energy technology guide. Paris: IEA; 2024. https://www.iea.org/data-and-statistics/data-tools/etp-clean-energy-technology-guide.
[55] Malhotra A, Schmidt TS. Accelerating low-carbon innovation. Joule 2020;4(11): 2259–67. https://doi.org/10.1016/j.joule.2020.09.004.